

ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2018/19

DOCTORANDO: **LLOSENT DE NARDIZ, MARÍA DEL DULCE NOMBRE**
D.N.I./PASAPORTE: ****9964R

PROGRAMA DE DOCTORADO: **D402 ESTUDIOS NORTEAMERICANOS**
DPTO. COORDINADOR DEL PROGRAMA: **INST.UNIV.INV. EST. NORTEAME. BENJ. FRANKLIN**
TITULACIÓN DE DOCTOR EN: **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

En el día de hoy 25/01/19, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de **ENRIQUE ALONSO GARCÍA**.

Sobre el siguiente tema: **LA INTEROPERABILIDAD LEGAL EN BIOINFORMÁTICA A PARTIR DE LA EXPERIENCIA NORTEAMERICANA Y LA EUROPEA: ¿ES POSIBLE LA INTEROPERABILIDAD LEGAL A NIVEL GLOBAL?**

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL³ de (no apto, aprobado, notable y sobresaliente): Sobresaliente

Alcalá de Henares, 25 de enero de 2019

EL PRESIDENTE



Fdo.:

ARTURO H. ARIÑO

EL SECRETARIO



Fdo.:

Julio César Somo

EL VOCAL



Fdo.:

Pablo G. Méxica

Con fecha 25 de Enero de 2019 la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- ☒ Conceder la Mención de "Cum Laude"
☐ No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada

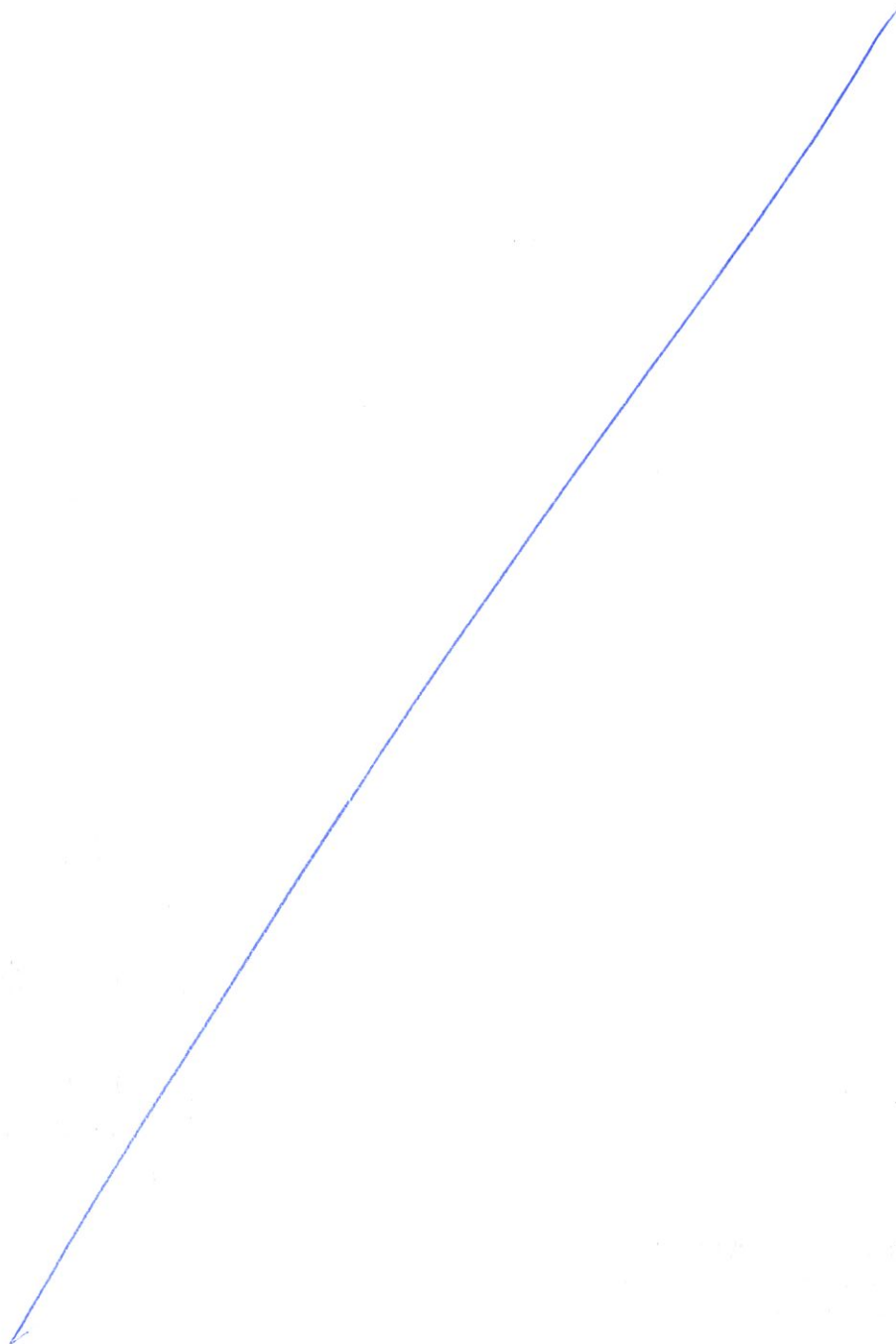


Fdo.: **LLOSENT DE NARDIZ, MARÍA DEL DULCE NOMBRE**



³ La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:



En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 25 de febrero, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por LLOSENT DE NARDIZ, MARÍA DEL DULCE NOMBRE, el día 25 de enero de 2019, titulada *LA INTEROPERABILIDAD LEGAL EN BIOINFORMÁTICA A PARTIR DE LA EXPERIENCIA NORTEAMERICANA Y LA EUROPEA: ¿ES POSIBLE LA INTEROPERABILIDAD LEGAL A NIVEL GLOBAL?*, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve otorgar** a dicha tesis la

MENTIÓN "CUM LAUDE"

Alcalá de Henares, 25 de febrero de 2019
EL VICERRECTOR DE INVESTIGACIÓN Y TRANSFERENCIA


Javier de la Mata de la Mata

Copia por e-mail a:

Doctorando: LLOSENT DE NARDIZ, MARÍA DEL DULCE NOMBRE

Secretario del Tribunal:

Director de Tesis: ENRIQUE ALONSO GARCÍA



Universidad
de Alcalá

ESCUELA DE DOCTORADO
Servicio de Estudios Oficiales de
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D^a _____
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo
presentado la misma en formato: ☐ soporte electrónico ☐ impreso en papel, para el depósito de la
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: _____ se procede, con
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a _____ de _____ de 20____



Fdo. El Funcionario



Carmen Flys Junquera, Coordinador de la Comisión Académica
del Programa de Doctorado en Estudios Norteamericanos

INFORMA que la Tesis Doctoral titulada La interoperabilidad legal en bioinformática a partir de la experiencia norteamericana y la europea: ¿Es posible la interoperabilidad legal a nivel global?, presentada por D/D^a María del Dulce Nombre Lloset de Nardiz, bajo la dirección del / de la Dr/a. Enrique Alonso García, reúne los requisitos científicos de originalidad y rigor metodológicos para ser defendida ante un tribunal. Esta Comisión ha tenido también en cuenta la evaluación positiva anual del doctorando, habiendo obtenido las correspondientes competencias establecidas en el Programa.

Para que así conste y surta los efectos oportunos, se firma el presente informe en Alcalá de Henares a 19 de junio de 2018

Fdo.:


Carmen Flys



Programa de doctorado en Doctorado en
Estudios Norteamericanos (RD 99/2011)

*La interoperabilidad legal en bioinformática a partir
de la experiencia norteamericana y la europea: ¿Es
posible la interoperabilidad legal a nivel global?*

Tesis doctoral presentada por:

María del Dulce Nombre Llorent de Nardiz

Director:

Dr. Enrique Alonso García

Alcalá de Henares, 2018

Informe favorable de la tesis



INSTITUTO UNIVERSITARIO
DE INVESTIGACIÓN EN
ESTUDIOS NORTEAMERICANOS
"BENJAMIN FRANKLIN"
C/ Trinidad, 1
28801 Alcalá de Henares (Madrid, Spain)
Teléfono: +34 91 885 5252
Fax: +34 91 885 5248
<http://www.institutofranklin.net>

El Profesor D. Enrique Alonso García, director de la Tesis Doctoral con el título "La interoperabilidad legal en bioinformática a partir de la experiencia norteamericana y la europea: ¿es posible la interoperabilidad legal a nivel global?" emite

INFORME FAVORABLE

Doña María del Dulce Nombre Lloset de Nárdiz ha realizado bajo mi dirección su Tesis Doctoral titulada "La interoperabilidad legal en bioinformática a partir de la experiencia norteamericana y la europea: ¿es posible la interoperabilidad legal a nivel global?", que está concluida y que reúne, a mi juicio, las condiciones necesarias exigidas para acceder al grado de Doctor.

La tesis está claramente estructurada y parte de una sólida base teórica, realizando un análisis de la interoperabilidad legal -es decir, las fórmulas para lograr reducir o eliminar las restricciones jurídicas a las que está sometido el flujo de datos digitales-, en un ámbito concreto: la bioinformática, un campo en el que confluyen muchas ciencias y distintos tipos de datos: científicos, económico-sociales e incluso de humanidades. Y lo hace comparando dos marcos normativos: el europeo y el estadounidense.

Para poder realizar dicho análisis se parte de la necesidad de contextualizarlo en sus dos componentes básicos: por un lado, la regulación de las nuevas tecnologías de la información (en particular las de carácter infraestructural) y los principios que regulan el flujo masivo de datos científicos; por otro lado, la regulación de la conservación del uso sostenible de la diversidad biológica.

El primer marco de contextualización lo centra la doctoranda en lo que han supuesto, para el derecho europeo y el norteamericano, la digitalización y la economía del conocimiento. Partiendo del análisis del nacimiento de Internet y los efectos que ello ha tenido, describe la aparición de nuevas formas de gestión del conocimiento: las e-infraestructuras y la utilización por las mismas de los datos digitales y su regulación: el concepto mismo de dato, de su propiedad, si es que son apropiables y su protección por parte de aquellos que los generan y organizan. Y, en particular, lo hace estudiando los nuevos sistemas (*open access*, *public access*, *open data*, *open science*, etc.) que exigen a los científicos poner al alcance de la ciencia en general y, por tanto de otros científicos y de las infraestructuras, los datos digitalizados derivados de sus investigaciones para evitar el efecto que sobre la



INSTITUTO UNIVERSITARIO
DE INVESTIGACIÓN EN
ESTUDIOS NORTEAMERICANOS
"BENJAMIN FRANKLIN"

C/ Trinidad, 1
28801 Alcalá de Henares (Madrid, Spain)
Teléfono: +34 91 885 5252
Fax: +34 91 885 5248
<http://www.institutofranklin.net>

ciencia estaba generando el exceso de "privatización" de los mismos a partir de un mal entendimiento de cuáles deben ser los derechos de propiedad intelectual sobre dichos datos científicos, en especial cuando hubieran sido obtenidos en proyectos de investigación financiados con fondos públicos; todo ello en el contexto de una nueva cultura de *data sharing*. Como ejemplo de esa nueva cultura y regulación del flujo de datos se examinan dos infraestructuras: GEOSS y CODATA. La interoperabilidad legal requiere dar acceso a los datos, idealmente, sin necesidad de intervención de abogados permitiendo su reproducibilidad, a partir de la implantación masiva de los llamados *FAIR principles* (los datos deben ser *findable, accesible, interoperable and reusable*). La construcción del concepto de interoperabilidad legal, pues, precisa de formas novedosas más allá de las regulaciones nacionales y la creación de instrumentos como pueden ser la dedicación de los datos al dominio público o el resto de las licencias Creative Commons, en particular la dedicación y licencias CC0 y CC-BY. También describe el trabajo de organizaciones internacionales que facilitan instrumentos, en cuya creación ha tomado parte la doctoranda, creados como instrumento de colaboración en la gestión de datos científicos por la Comisión Europea (en concreto la DG Research) y el Gobierno norteamericano (a través de la NSF –National Science Foundation), en concreto, la Research Data Alliance y sus Principios y Directrices.

En el segundo marco de contextualización, la doctoranda analiza el derecho global de la diversidad biológica que se inicia a principios de los años 70 y se plasma en la Convención para la Diversidad Biológica de 1992. La doctoranda analiza, también, la Plataforma Intergubernamental sobre Biodiversidad y Servicios de los Ecosistemas (y PBES) como el instrumento que trata de recoger el testigo de la CDB cuyo órgano subsidiario científico, creado por la Convención, no acabó de funcionar correctamente, siendo sustituido, o mejor dicho, completado, por dicha Plataforma.

Después de esta doble contextualización, en la segunda parte de la tesis, la doctoranda analiza, en profundidad, el impacto de la aplicación de las nuevas tecnologías a la conservación de la biodiversidad; es decir, la confluencia de ambos contextos en la creación de instrumentos de bioinformática (la aplicación de la informática a la biodiversidad).

Aborda, para ello, el análisis de la creación de la Global Biodiversity Information Facility (GBIF) y, sobre todo, de la elaboración y publicación del *Global Biodiversity Informatics Outlook (GBIO): Delivering Biodiversity knowledge in the information age*, en el marco de la reunión de los 100 expertos de bioinformática, en la que participó la doctoranda.

Este análisis se ha llevado a cabo mediante la participación de la misma en dos proyectos europeos dentro del marco de los Programas FP7 y H2020. Se trata en concreto de dos



INSTITUTO UNIVERSITARIO
DE INVESTIGACIÓN EN
ESTUDIOS NORTEAMERICANOS
"BENJAMIN FRANKLIN"
C/ Trinidad, 1
28801 Alcalá de Henares (Madrid, Spain)
Teléfono: +34 91 885 5252
Fax: +34 91 885 5248
<http://www.institutofranklin.net>

proyectos europeos que ha permitido estudiar en detalle, de qué manera se están aplicando todas las cuestiones teóricas a la realidad y formas de abordar la interoperabilidad legal en el funcionamiento de las infraestructuras de biodiversidad.

Antes de analizar el funcionamiento de dichas infraestructuras se estudia también GEO BON como instrumento de observación de biodiversidad enmarcado en el Grupo de Observación de la tierra (GEO).

Volviendo a las infraestructuras la tesis se adentra en la descripción de los modelos continentales de Estados Unidos (DataONE y NEON) y de Europa (LifeWatch), y algunos otros ejemplos como el modelo australiano (ALAS), brasileño (CRIA), o de la academia de ciencias de la República Popular China (CAS). Después del análisis de los distintos modelos continentales, la tesis aborda más en concreto los dos casos de estudio realizados en el marco de proyectos de la Unión Europea que sistematizan las distintas cuestiones estudiadas en la tesis: CREaTIVE-B y GLOBIS-B, proyectos en los que han colaborado las instituciones continentales invitadas, así como las propiamente globales (GEBIF y GEO BON), y en los que el Instituto Franklin de la Universidad de Alcalá llevó el liderazgo del Working Package IV, relativo a las cuestiones de Law & Policy, en muchas de cuyas discusiones participó la doctoranda.

Precisamente el segundo de los proyectos acaba de finalizar el 31 de mayo de 2018, momento en el cual también se cierra el análisis con la incorporación de las conclusiones finales como anexo de la tesis.

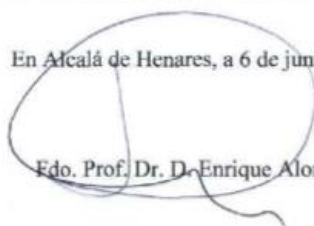
La amplia y actualizada bibliografía demuestra un amplio conocimiento teórico y práctico de la materia.

Por todo ello, considero que la tesis reúne los requisitos necesarios para ser defendida.

Y para que así conste, se firma el presente informe en la ciudad de Alcalá de Henares.

D. Enrique Alonso García

En Alcalá de Henares, a 6 de junio de 2018


Edo. Prof. Dr. D. Enrique Alonso García

Julio Cañero Serrano
Director
Instituto Franklin
Universidad de Alcalá

Certificados de presentación de la tesis



INSTITUTO UNIVERSITARIO
DE INVESTIGACIÓN EN
ESTUDIOS NORTEAMERICANOS
"BENJAMIN FRANKLIN"
C/ Trinidad, 1
28801 Alcalá de Henares (Madrid, Spain)
Teléfono: +34 91 885 5252
Fax: +34 91 885 5248
<http://www.institutofranklin.net>

Alcalá de Henares, 15 de mayo de 2018

D. JULIO CAÑERO SERRANO, PROFESOR TITULAR DE UNIVERSIDAD DE FILOLOGÍA INGLESA, Y DIRECTOR DEL INSTITUTO UNIVERSITARIO DE INVESTIGACIÓN EN ESTUDIOS NORTEAMERICANOS "BENJAMIN FRANKLIN" DE LA UNIVERSIDAD DE ALCALÁ,

INFORMA:

Que, a la vista del informe favorable del Profesor doctor Enrique Alonso sobre la tesis doctoral realizada en este Instituto, y bajo su dirección, por D^a. María del Dulce Nombre Lloset de Nárdiz, y titulada "LA INTEROPERABILIDAD LEGAL EN BIOINFORMÁTICA A PARTIR DE LA EXPERIENCIA NORTEAMERICANA Y LA EUROPEA: ¿ES POSIBLE LA INTEROPERABILIDAD LEGAL A NIVEL GLOBAL?," se autoriza por parte de este Instituto su presentación para que se proceda a su ulterior tramitación y lectura pública.


Julio Cañero Serrano
Director
Instituto Franklin
Universidad de Alcalá



El futuro tiene muchos nombres. Para los débiles es lo inalcanzable. Para los temerosos, lo desconocido. Para los valientes es la oportunidad.

VÍCTOR HUGO

Ten el coraje de seguir a tu corazón y tu intuición, ellos ya saben de algún modo lo que quieres llegar a ser.

STEVE JOBS

Quiero dedicar mi tesis a mi padre. Cuando terminé mi carrera, me dijo que el futuro estaba en el medioambiente y en la tecnología, que tenía que formarme en esos dos ámbitos del derecho. Y, de alguna manera, aquí estoy...

Aunque ya no está aquí, me ha ido guiando y, por una serie de causalidades, no creo mucho en la casualidad, he terminado escribiendo mi tesis doctoral en un tema donde se reúnen las dos grandes cuestiones «del futuro» que él me decía: la tecnología y la biodiversidad. Su recuerdo me ha acompañado mucho siempre y, sobre todo, durante este camino.

También quiero dedicarle la tesis a Natalia; sin su ayuda y apoyo esto no hubiera sido posible. Ha creído en mí..., y me ha ayudado a salir en los momentos en los que me he atascado. Gracias a eso, he sido capaz de llegar hasta aquí y, sin lugar a dudas, es la que más me ha sufrido en el proceso, desde el principio hasta el final. Y claro a Alejandra, que ya está en casa...

Agradecimientos

Desde que terminé la carrera, quise hacer el doctorado para poder dedicarme, siquiera en parte, a la enseñanza en la Universidad. Durante años no fue posible realizarlo por diversas razones que ya no vienen al caso.

En 2010, tuve la suerte de conocer al Excmo. Sr. D. Enrique Alonso García, quien me propuso escribir mi tesis sobre temas relacionados con la bioinformática o, más correctamente, con la biodiversidad informática. En mi carrera había trabajado en la industria farmacéutica trabajando con toda la problemática relacionada con la propiedad intelectual e industrial, con cuestiones relacionadas con la protección de datos, etc.

Él vio más allá y confió en mí invitándome a formar parte de los distintos proyectos en el marco de los cuales he desarrollado mi tesis. Gracias, por el cariño, por la confianza. He aprendido mucho a su lado, tanto de forma (como persona) como de fondo (como jurista), y me he quedado con ganas de más.

Muchas gracias al Instituto de Estudios Norteamericanos Franklin de la Universidad de Alcalá, en el cual se desarrollan estas iniciativas de intercambio entre la Unión Europea y Estados Unidos que permiten hacer un análisis comparado de ambos sistemas y ver formas de colaboración, y a D. José Antonio Gurpegui Palacios, por su apoyo en todo momento, desde la Universidad de Alcalá, de forma directa e indirecta. Más en concreto, al programa de «Friends of Thoreau» del Instituto y a su coordinadora, Ana Recarte, que ha servido de paraguas para esta tesis por los medios que ha puesto a mi disposición de consulta, sin los cuales no hubiera sido posible desarrollar esta investigación, y especialmente por haberme incluido como parte del equipo de los proyectos de la DG Research de la Comisión Europea CREaTIVE-B con el que empezó todo el proyecto y posteriormente el GLOBIS-B del 7 Programa Marco y del H2020, de Investigación, respectivamente, de cuyos *Working Packages* números 4 respectivos ha estado a cargo el citado programa «Friends of Thoreau», en colaboración con múltiples universidades y centros de investigación europeos y norteamericanos.

A mi madre, quien, aunque no siempre me entiende, está como sabe y puede.

A mis hermanos, mis incondicionales a los que adoro y con los que he disfrutado y aprendido.

A mis abuelas, Pitis y Lola; por alguna razón las tengo muy presentes en mi vida aunque ya no están conmigo. Cada una a su manera me han acompañado y ayudado en mi vida. Sin ellas, hoy no sería yo.

Y a toda mi familia que, de una forma o de otra, cada uno, como sabe está y me ayuda.

María Llorent de Nardiz

Barcelona 10 de mayo de 2018

ÍNDICE

ÍNDICE DE ABREVIATURAS	XXI
INTRODUCCIÓN	25
1. JUSTIFICACIÓN DEL TEMA DE ESTUDIO	25
2. ESTRUCTURA DE LA INVESTIGACIÓN	27
2.1. DOS MARCOS DE CONTEXTUALIZACIÓN	27
A. La revolución digital	27
A.1. La aparición de un nuevo paradigma: Internet	27
A.2. La gestión del conocimiento	28
A.3. Regulación de las e-infraestructuras: diferencias entre la Unión europea y Estados Unidos	29
A.4. Las cuestiones jurídicas que plantea la revolución digital	30
A.5. Data sharing	34
B. La biodiversidad	36
2.2. LA APLICACIÓN DE LAS NUEVAS TECNOLOGÍAS SURGIDAS DE LA REVOLUCIÓN DIGITAL A LA CIENCIA DE LA BIODIVERSIDAD: LA BIOINFORMÁTICA	41
BLOQUE I. MARCOS PREVIOS DE CONTEXTUALIZACIÓN	49
CAPÍTULO 1. PRIMER MARCO DE CONTEXTUALIZACIÓN: LA REVOLUCIÓN DIGITAL	51
1. LA ECONOMÍA DEL CONOCIMIENTO: LA REVOLUCIÓN DIGITAL (DESDE MEDIADOS DEL SIGLO XX HASTA LA ACTUALIDAD)	51
1.1. LOS ORÍGENES DE INTERNET	51
1.2. EL CONCEPTO DE INTERNET	56
1.3. LA GESTIÓN DEL CONOCIMIENTO EN LA ECONOMÍA DEL CONOCIMIENTO	59
1.4. DEL <i>DATA WAREHOUSE</i> AL <i>BIG DATA</i>	64
2. LA NUEVA FORMA DE GESTIONAR EL CONOCIMIENTO EN EL SIGLO XXI: ¿QUÉ SON LAS E-INFRAESTRUCTURAS?	68
2.1. LAS FUNCIONES DE LAS CIBERINFRAESTRUCTURAS O E-INFRAESTRUCTURAS	71
2.2. DE LA CIENCIA TRADICIONAL A LA E-CIENCIA	74
2.3. UN ESTUDIO COMPARADO DE DOS ENFOQUES DISTINTOS EN LA REGULACIÓN DE LAS E-INFRAESTRUCTURAS	79
2.3.1. El enfoque en Europa: marco regulatorio, la creación de nuevas formas jurídicas	79
2.3.1.1. Análisis de los European Strategy Forum on Research Infrastructures (ESFRI)	79
2.3.1.2. Los Consorcios de Infraestructuras de Investigación Europeas	82

2.3.2. El enfoque en Estados Unidos: las políticas estadounidenses	89
3. LAS CUESTIONES JURÍDICAS QUE PLANTEA LA REVOLUCIÓN DIGITAL	90
3.1. UNA PRIMERA REFLEXIÓN SOBRE LAS CUESTIONES JURÍDICAS PREVIAS A LA INTEROPERABILIDAD	90
3.1.1. El concepto de dato: su estatuto jurídico	90
3.1.2. El concepto de dato científico (research data)	92
3.1.3. Los obstáculos jurídicos en relación con el concepto de dato	98
3.1.3.1. La propiedad de los datos	98
3.1.3.2. La protección de los datos	102
3.2. DOS ENFOQUES DISTINTOS: EUROPA FRENTE A ESTADOS UNIDOS	106
3.2.1. La protección de las bases de datos	106
3.2.1.1. El modelo europeo	107
A. La protección del contenido (los datos)	109
B. La protección de la estructura (el continente)	113
C. Los derechos <i>sui generis</i>	114
3.2.1.2. El modelo estadounidense (US Copyright Office, 1997)	118
A. Los datos generados por agencias gubernamentales	123
B. Los datos generados por instituciones académicas e instituciones sin ánimo de lucro	126
C. El sector privado	127
3.2.1.3. Un análisis comparado entre la protección de las bases de datos en Europa y en Estados Unidos: conclusiones	128
3.2.1.4. Una breve reflexión sobre los datos de biodiversidad	135
3.3. LAS TENDENCIAS <i>OPEN</i> (<i>OPEN DATA</i> Y <i>OPEN ACCESS</i> O ACCESO ABIERTO): EL NUEVO MARCO PARA EL INTERCAMBIO DE DATOS	136
3.3.1. Las distintas modalidades de open	137
3.3.1.1. <i>Open access</i> o acceso abierto (OA)	137
3.3.1.2. <i>Open data</i>	142
3.3.1.3. <i>Open source</i> o <i>software</i> con código abierto	144
3.3.2. El modelo europeo	148
3.3.3. El marco legislativo y las iniciativas políticas	151
3.3.3.1. La directiva sobre la reutilización de la información del sector público	151
3.3.3.2. La Directiva INSPIRE	155
3.3.3.3. El mercado único digital	162
3.3.3.4. La Agenda Digital para Europa	164
3.3.3.5. La licencia sobre el <i>text and data mining</i> : ¿un paso atrás?	165
3.3.4. El modelo estadounidense: las políticas de public access	170
3.3.4.1. Las exigencias de la FASTR	172
3.3.4.2. <i>Public access</i> en la Administración Obama	173
3.3.4.3. La Administración Trump	174
3.3.4.4. ¿Una excepción al <i>public access</i> ?: la University and Small Business Patent Procedures Act más conocida con el nombre de ley Bayh-Dole	176
4. EL <i>DATA SHARING</i> O INTERCAMBIO DE DATOS	181
4.1. EL SISTEMA DE SISTEMAS GLOBAL DE OBSERVACIÓN DE LA TIERRA	186
4.1.1. Los antecedentes del GEOSS: GEO	186
4.1.2. El GEOSS	188
4.1.2.1. El objetivo del GEOSS	191

4.1.2.2. Los <i>data sharing principles</i>	194
4.1.2.3. Los <i>data management principles</i>	200
4.2. COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY (CODATA). <i>POLICIES FOR OPEN ACCESS TO RESEARCH DATA</i>	202
4.3. LA INTEROPERABILIDAD LEGAL	206
4.3.1. Las posibles alternativas al copyright: el nacimiento del movimiento copyleft	216
4.3.2. Ejemplos concretos del movimiento copyleft: de las GNU a las licencias Creative Commons	216
4.3.3. Distintas propuestas para ir construyendo una interoperabilidad legal ¿global?	220
4.3.3.1. Los principios y directrices de implementación de la Research Data Alliance	220
4.3.3.2. Los principios del Foro de Belmont	225
4.3.3.3. Los <i>Principles for open data in science</i> o Principios Panton	228
4.3.3.4. The FAIR Guiding Data Principles	231
4.4. DOMINIO PÚBLICO: ¿ALTERNATIVA AL <i>COPYLEFT</i> O A LAS LICENCIAS CREATIVE COMMONS?	
¿PRIMEROS PASOS HACIA LA INTEROPERABILIDAD GLOBAL?	235
4.5. ¿CENTRALIZACIÓN DE LOS DATOS?	238

CAPÍTULO 2. SEGUNDO MARCO DE CONTEXTUALIZACIÓN: LA DIVERSIDAD BIOLÓGICA, TAMBIÉN LLAMADA BIODIVERSIDAD

1. LA CONCIENCIACIÓN EN MATERIA DE BIODIVERSIDAD: ORÍGENES DEL TÉRMINO	241
2. EL CONVENIO SOBRE LA DIVERSIDAD BIOLÓGICA (CDB)	244
2.1. LA GESTACIÓN DEL CDB	244
2.2. ESTADOS UNIDOS: SU NEGATIVA A RATIFICAR EL CDB	247
2.3. LOS ÓRGANOS DE GOBIERNO DEL CDB	250
2.3.1. La Conferencia de las Partes (CoP)	251
2.3.2. La Secretaría Ejecutiva	253
2.3.3. Órgano Subsidiario de Asesoramiento Científico, Técnico y Tecnológico	254
2.4. ¿EL FRACASO DEL SBSTTA?	257
2.5. EL CDB Y LOS OBJETIVOS DE FRENAR LA PÉRDIDA DE DIVERSIDAD BIOLÓGICA	259
3. LA PLATAFORMA INTERGUBERNAMENTAL SOBRE BIODIVERSIDAD Y SERVICIOS DE LOS ECOSISTEMAS	264
3.1. LOS ORÍGENES DE IPBES	265
3.2. EL NACIMIENTO DE IPBES	269
3.3. LOS RETOS DEL IPBES	272
3.4. LAS DIFERENCIAS Y SIMILITUDES ENTRE IPBES E IPCC	274

BLOQUE II. LA APLICACIÓN DE LAS NUEVAS TECNOLOGÍAS A LA BIODIVERSIDAD: EL NACIMIENTO DE LA INFORMÁTICA APLICADA A LA BIODIVERSIDAD O BIOINFORMÁTICA

CAPÍTULO 1. EL NACIMIENTO DEL TÉRMINO BIOINFORMÁTICA

CAPÍTULO 2. LOS PRIMEROS PASOS DE LA BIOINFORMÁTICA: LA CREACIÓN DE LA INFRAESTRUCTURA MUNDIAL DE INFORMACIÓN EN BIODIVERSIDAD

1. LA CONFIGURACIÓN DEL GBIF	285
2. LA CREACIÓN DEL GBIF	285

3. LA MISIÓN DEL GBIF	287
3.1. LA RELACIÓN ENTRE EL GBIF Y EL IPBES	288
3.2. EL GBIF: UNA HERRAMIENTA GLOBAL DE LA BIOINFORMÁTICA	291
3.3. EL GBIF DATA SHARING	293
4. EL GLOBAL BIODIVERSITY INFORMATICS OUTLOOK (GBIO): DELIVERING BIODIVERSITY KNOWLEDGE IN THE INFORMATION AGE	296
4.1. EL ORIGEN DEL GBIO (GBIF, 2015)	296
4.2. LAS CONCLUSIONES DEL GBIO	297
4.2.1.- El área de cultura	300
4.2.2.- El área de datos	301
4.2.3.- El área de evidencias	303
4.2.4.- El área de la comprensión	304
<u>CAPÍTULO 3. EL GRUPO DE OBSERVACIÓN DE LA TIERRA EN BIODIVERSIDAD</u>	<u>307</u>
1. ANTECEDENTES	307
2. ESTRUCTURA	312
3. OBJETIVO	313
4. EL OBSERVATORIO EUROPEO DE BIODIVERSIDAD	316
<u>CAPÍTULO 4. LA APLICACIÓN EN UNA ESCALA INFERIOR A LA GLOBAL, DOS MODELOS CONTINENTALES DE GESTIÓN DE LA BIOINFORMÁTICA: EUROPA Y ESTADOS UNIDOS. OTROS EJEMPLOS DE MODELOS CONTINENTALES</u>	<u>319</u>
1. EL MODELO EUROPEO: LIFEWATCH ERIC (E-SCIENCE AND TECHNOLOGY INFRASTRUCTURE FOR BIODIVERSITY DATA AND OBSERVATORIES)	319
2. DOS EJEMPLOS DE ESTADOS UNIDOS	328
2.1. EL MODELO DE DATA OBSERVATION NETWORK FOR EARTH (DATAONE)	328
2.1.1. La estructura de DataONE	329
2.2. EL EJEMPLO DE LA NATIONAL ECOLOGICAL OBSERVATORY NETWORK (NEON)	336
2.2.1. La estructura de NEON	337
3. OTROS MODELOS CONTINENTALES	341
3.1. SOUTH AFRICAN NATIONAL BIODIVERSITY INSTITUTE	341
3.2. CENTRO DE REFERENCIA EM INFORMAÇÃO AMBIENTAL (CRIA), BRASIL	344
3.3. THE ATLAS OF LIVING AUSTRALIA (ALA)	345
3.4. ACADEMIA DE CIENCIAS DE CHINA (CAS CHINESE ACADEMY OF SCIENCE)	350
<u>CAPÍTULO 5. COORDINACIÓN PARA PODER REALIZAR ANÁLISIS GLOBALES EN LA ESCALA ESPACIAL Y TEMPORAL</u>	<u>351</u>
1. LA CONSTRUCCIÓN DE ROADMAPS PARA FAVORECER EL TRABAJO CONJUNTO: LOS PROGRAMAS DE PLANIFICACIÓN Y MODELIZACIÓN	354
1.1. UNA APLICACIÓN PRÁCTICA DE TODO LO ANTERIOR A UN PROBLEMA Y CÓMO SE HA TRABAJADO EN EL PROYECTO CREATIVE-B	354
1.2. EL ROADMAP	357
1.2.1. Un paso más, GLOBIS-B	362
1.3. LAS VARIABLES ESENCIALES DE BIODIVERSIDAD O EBV	366

2. LAS CUESTIONES PRÁCTICAS QUE PLANTEA LA INTEROPERABILIDAD LEGAL EN EL ÁMBITO DE LOS DATOS DE BIODIVERSIDAD	377
2.1. LA INTEROPERABILIDAD LEGAL EN EL ÁMBITO DEL CREATIVE-B	377
2.2. LA INTEROPERABILIDAD LEGAL EN EL ÁMBITO DEL GLOBIS-B	378
CONCLUSIONES	385
DOS MARCOS DE CONTEXTUALIZACIÓN	385
A. PRIMER MARCO CONTEXTUAL: LA REVOLUCIÓN DIGITAL	385
A.1. La aparición de las nuevas tecnologías y la creación de un nuevo contexto: Internet	385
A.2. La revolución del conocimiento.	386
A.3. Nuevas formas de gestión del conocimiento: las infraestructuras del conocimiento	387
A.3.1. Las infraestructuras del conocimiento	388
A.3.2. La aparición de la e-ciencia	389
A.4. Dos enfoques diferentes: e-infraestructuras o ciber-infraestructuras	389
A.4.1. El sistema europeo: la e-infraestructuras	389
A.4.2. El sistema de los Estados Unidos: las ciber-infraestructuras.	392
A.5. Las implicaciones de la revolución digital: la interoperabilidad	393
A.6. La apropiabilidad de los datos	395
A.7. La protección de los datos mediante sistemas de propiedad intelectual	396
A.7.1. El sistema europeo de protección	397
A.7.2. El sistema de protección de Estados Unidos	399
A.7.3. Conclusiones del análisis de ambos sistemas	400
A.8. La aparición de un nuevo marco de intercambio de datos: las tendencias open	401
A.8.1. La construcción del <i>open access</i> en Europa	405
A.8.2. Estados Unidos y las políticas de <i>public access</i>	407
A.9. El data sharing o intercambio de datos	411
A.10. La interoperabilidad legal	412
A.11. El dominio público el movimiento copyleft y las licencias Creative Commons: primeros pasos hacia la interoperabilidad global	415
B. SEGUNDO MARCO CONTEXTUAL: LA BIODIVERSIDAD	416
B.1. El concepto de biodiversidad	416
B.2. El CDB	417
B.3. El IPBES	421
LA APLICACIÓN DE LAS NUEVAS TECNOLOGÍAS DIGITALES A LA BIODIVERSIDAD: EL NACIMIENTO DE A INFORMÁTICA APLICADA A LA BIODIVERSIDAD O BIOINFORMÁTICA	422
A. EL NACIMIENTO DEL TÉRMINO BIOINFORMÁTICA	422
B. PRIMEROS PASOS DE LA BIOINFORMÁTICA: LA CREACIÓN DEL GBIF	422
B.1. El GBIO	423
C. GEO BON	425
D. APLICACIÓN DE UNA ESCALA INFERIOR A LA GLOBAL: DOS MODELOS CONTINENTALES DE GESTIÓN DE LA BIOINFORMÁTICA	428
D.1. El modelo europeo: LifeWatch	428
E. DOS EJEMPLOS DE ESTADOS UNIDOS	429
E.1. Data ONE	429
E.2. NEON	430

F. OTROS MODELOS CONTINENTALES	430
G. COORDINACIÓN PARA LA REALIZACIÓN DE ANÁLISIS GLOBALES EN LA ESCALA ESPACIAL Y TEMPORAL	433
G.1. CREaTIVE-B	433
G.2. GLOBIS-B	434
BIBLIOGRAFÍA	463
NORMAS Y RECOMENDACIONES	507
ESTADOS UNIDOS	507
Normas	507
Jurisprudencia	507
Informes	508
EUROPA	509
Normas	509
Informes	511
INTERNACIONAL	512
Acuerdos/tratados	512
Informes	513
ÍNDICE DE FIGURAS	517

Índice de abreviaturas

ABS	Access to genetic resources and benefitsharing
AIC	Acciones de Investigación Colaborativa
ALA	Atlas of Living Australia
ARPA	Agencia de Proyectos de Investigación Avanzados
BBMRI	Biobanking and Biomolecular resources Research Infrastructure
BHL	Biodiversity Heritage Library, Biblioteca sobre el Patrimonio de la Biodiversidad
BIP	Biodiversity Indicators Partnership
BOAI	Open Initiative de Budapest
BON	Biodiversity Observation Network
CAS	Academia China de Ciencias
CBOL	Consortium for the Barcode of Life, Consorcio del Código de Barras de la Vida
CC	Creative Commons
CDB	The Biodiversity convention, Convenio de Naciones Unidas de Diversidad Biológica
CEI	Consejo Europeo de Investigación
CERIC	Central European Research Infrastructure
CERN	Conseil Européen pour la Recherche Nucléaire, Centro Europeo de Investigaciones Nucleares
CESSDA	Council of European Social Science Data Archives
CITES	Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres
CLARIN	Common Language Resources and Technology Infrastructure
CLD	Convención de Lucha contra la Desertificación
CONTU	National Commission on New Technological Uses of Copyrighted Works
CoP	Conference of Parts, Conferencia de las Partes

CODATA	Committee on Data for Science and Technology, Comité de Información para Ciencia y Tecnología
CReACTIVE-B	Coordination of Research e-Infrastructures Activities Toward an International Virtual Environment for Biodiversity
CRIA	Centro de Referencia en Informação Ambiental
CRM	customer relationship management
DADI	Data Access and Data Interoperability
Data-COR	Data Collection of Open Resources for Everyone
Data ONE	Data Observation Network for Earth
DIGIT	Digitization of Natural History Collection Data
DMCA	Digital Millennium Copyright Act
DOCE	Diario Oficial de las Comunidades Europeas
DoD	Departamento de Defensa americano
DTL	Dutch Techcentre for Life Sciences, Centro Tecnológico Holandés de Ciencias de la Vida
EATRIS	European infrastructure for translational medicine
EBV	essential biodiversity variables
ECAT	Electronic Catalog of Names of Known Organisms
ECRIN	European Clinical Research Infrastructure Network
ECSEL	European Carbon Dioxide Capture and Storage Laboratory Infrastructure
EMA	European Medicines Agency, Agencia Europea del Medicamento
EMBL	European Molecular Biology Laboratory
EMSO	European Multidisciplinary Seafloor and Water Column Observatory
EOL	Encyclopedia of Life, Enciclopedia de la Vida
ERIC	European Research Infrastructure Consortium, Consorcios de Infraestructuras de Investigación Europeas
ESFRI	European Strategy Forum on Research Infrastructures, Foro Estratégico Europeo sobre Infraestructuras de Investigación
ESIP	Earth Science Information Partners
ESS	European Spallation Source-ERIC
ESS	European Social Survey
EU BON	European Biodiversity Observation Network

EUDAT	European Data Infrastructure
FAIR	findable, accesible, interoperable y reusable.
FAPESP	Fundación de Investigación de Sao Paulo
FASTR	Fair Access to Science and Technology Research
GBIC	Global Biodiversity Informatics Conference
GBIF	Global Biodiversity Informatics Facility, Infraestructura Mundial de Información en Biodiversidad
GBIO	Global Biodiversity Informatics Outlook
GEOSS	Global Earth Observation System of Systems
GEO	Grupo sobre Observaciones Terrestres
GEO BON	Group on Earth Observations on Biodiversity Observation, Grupo de Observación de la Tierra de Observación de la Biodiversidad
GEO CODATA	Comité de Datos sobre Ciencia y Tecnología del ICSU
GEOSS	Global Earth Observation System of Systems, Sistema Global de Sistemas de Observación de la Tierra
GBIC	Global Biodiversity Informatics Conference
ICSU	International Social Science Council, Consejo Internacional de Ciencia
ICOS	Integrated Carbon Observation System, Sistema Integrado de Observación del Carbono
IIASA	Instituto Internacional de Análisis de Sistemas Aplicados
IMoSEB	International Mechanism of Scientific Expertise on Biodiversity
	International Treaty on Plant Genetic Resources for Food and Agriculture
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Plataforma Intergubernamental sobre Biodiversidad y Servicios de los Ecosistemas
IPCC	Intergovernmental Panel on Climate Change. Panel de Cambio Climático
IPGRI:	Instituto Internacional de Recursos Fitogenéticos
ITER	International Thermonuclear Experimental Reactor
ISSC	Consejo Internacional de Ciencias Sociales
JIV	Joint Institute for VLBI
MEA	Millenium Ecosystem Assessment, Evaluación de los Ecosistemas del Milenio

NASA	National Aeronautics and Space Administration
NBII	National Biological Information Infrastructure
NCBI	National Center for Biotechnology Information, Centro Nacional de Información Biotecnológica
NCRIS	National Collaborative Research Infrastructure Strategy, Estrategia de Infraestructuras de Investigación Colaborativas
NIH	National Institute of Health
NEON	National Ecological Observatory Network
OA	open access
OCB	Outreach and Capacity Building
OKF	Open Knowledge Foundation, Fundación de Conocimiento Abierto
OMB	Office of Management and Budget, Oficina de Gestión y Presupuestos
OpenAIRE	Open Access Infrastructure for Research in Europe
PNUMA	Programa de las Naciones Unidas para el Medio Ambiente
RDA	Research Data Alliance
SANBI	South African Biodiversity Institute
SBSTTA	Subsidiary Body on Scientific, Technological and Technical Advice, Órgano Subsidiario de Asesoramiento Científico, Técnico y Tecnológico
SHARE	Survey of Health, Ageing and Retirement in Europe
TDM	text and data mining<
TDWG	Biodiversity Information Standards, Estándares de Información para la Biodiversidad
UE	Unión Europea
UNFCCC	United Nations Framework Convention on Climate Change Convención, Marco de las Naciones Unidas sobre el Cambio Climático
USGCRP	United States Global Change Research Program, Programa de Investigación sobre Cambio Global de Estados Unidos
USGS	United States Geological Survey
VRE	virtual research environments
WFCC	World Federation of Culture Collections
WORMS	World Register of Marine Species

Introducción

El objetivo de esta tesis es triple. Por una parte, se aborda el estudio de dos marcos contextuales previos para poder analizar idealmente el tercero, que constituye el núcleo de la tesis. El primero de ellos consiste en el análisis del cambio de paradigma que ha traído la revolución digital y las cuestiones jurídicas que se plantean en relación con la accesibilidad, interoperabilidad y reutilización de los datos digitales, analizando lo que se conoce con el nombre de interoperabilidad legal. El segundo es el análisis de la construcción de la ciencia de la diversidad biológica (en adelante, biodiversidad). El tercero aborda la construcción conceptual de una nueva disciplina: la bioinformática: la aplicación de la informática a la biodiversidad, es decir, el resultado de aplicar a la ciencia y a los datos relacionados con la biodiversidad aquella revolución digital.

Así pues, como consecuencia lógica de la fusión de ambos conceptos, la tesis trata de los problemas jurídicos que plantea la bioinformática como área de conocimiento en la que confluyen ambos mundos conceptuales: por un lado, la biodiversidad como ciencia que, como todas las ciencias, tiene una historia larga basada en metodologías ajenas a la regulación de las tecnologías de revolución digital y, por otro, los cambios operados en la misma cuando su contenido da un salto cualitativo, es decir, precisamente cuando una de las metodologías consista precisamente en su digitalización o, si se prefiere, en la digitalización de los resultados de la investigación y en la aplicación de las propuestas de construcción y uso sostenible de la biodiversidad.

Conviene anunciar ya que la manera de abordar este mundo conceptual simbiótico y su regulación no se trata de la misma forma en todos los sistemas jurídicos. De hecho, se hará un análisis comparado de dos enfoques bien diferenciados: el de la Unión Europea y el de Estados Unidos.

1. JUSTIFICACIÓN DEL TEMA DE ESTUDIO

Lo que en la actualidad caracteriza la generación de conocimiento es tanto la velocidad de su transmisión como la extensión de su alcance geográfico y su volumen (tiempo y espacio). Los datos resultantes de un estudio pueden servir de base para un estudio diferente, pueden ser fuente generadora de nuevo

conocimiento. Si bien es cierto que esto ha existido siempre, lo ha hecho a una escala en la que la variable espacial era muy reducida y la variable temporal, muy larga. En cambio, ahora, el espacio en su doble o triple dimensión (extensión geográfica y volumen) es global y el tiempo es cuestión de segundos. Se hace, pues, necesario buscar nuevas soluciones a la gestión del conocimiento.

Sin embargo, la multiplicidad de orígenes (geográficos, idiomáticos, distintos derechos involucrados, sistemas, etc.) hace que el acceso y la reutilización de estos datos puedan quedar sujetas a restricciones y barreras cuyo origen y causa nada tienen que ver con la posibilidad real y técnica de obtener y procesar en segundos todos los conocimientos acumulados desde los orígenes de la historia hasta lo publicado hace segundos en lugares muy remotos. Es decir, lo que técnicamente es posible puede verse, y de hecho se ve, imposibilitado por el Derecho.

Con la finalidad de evitar que aparezcan barreras o de poner fin a las que puedan existir, surge la necesidad de construir la interoperabilidad de los sistemas. Existen barreras de distinta índole: técnica (estándares, compatibilidad de sistemas), interoperabilidad sintáctica y semántica (utilización de un lenguaje y con un mundo conceptual idéntico por todos los datos que deben intercambiarse), y también existen barreras jurídicas. Este trabajo, sin dejar de hacer una breve referencia a las primeras y a las segundas, dará más importancia a las terceras, es decir, en las jurídicas o, mejor dicho, en cómo eliminar dichas barreras jurídicas para lograr lo que, recientemente, se conoce con el nombre de «interoperabilidad legal».

Efectivamente, el concepto de «interoperabilidad legal» es una traducción literal del inglés (*legal interoperability*), ya que en castellano sería más correcto hablar de «interoperabilidad jurídica». Este concepto se entiende hoy universalmente (aunque desde hace poco tiempo) —desde una perspectiva amplia— como el lado jurídico o los aspectos jurídicos que necesitan ser tenidos en cuenta y solventados siempre que se pretenden lograr esquemas de interoperabilidad técnica, sintáctica y semántica de datos digitales, que la perspectiva legal siempre presupone, ya que opera a partir de aquellas o como un plus a las mismas (Alonso García, 2016b).

Aunque existe una multiplicidad de normas y regulaciones que afectan al acceso, intercambio y reutilización de los datos, cada vez aparecen más sistemas con un ámbito territorial amplio e incluso global. Por eso, la principal pregunta a responder es la siguiente: ¿es posible una interoperabilidad legal global, de manera que estos dos enfoques, el europeo y el estadounidense en particular, y otros que coexisten en paralelo converjan hacia una solución única?

Analizar la interoperabilidad legal ha sido la clave y el objeto principal de esta tesis, es decir, el análisis de cuáles son las fórmulas para lograr reducir o eliminar las restricciones jurídicas a las que está sometido el flujo de los datos digitales.

Para ello, se ha elegido un campo concreto en el que se ha construido una nueva disciplina: la bioinformática. Y no se ha hecho por azar, sino porque se trata de un campo en el que confluyen muchas ciencias y distintos tipos de datos (los puramente científicos con los económico-sociales e incluso los de las humanidades) con la finalidad de objetivar y facilitar la efectividad de una política pública que todos los Estados del mundo se han comprometido a implantar y mantener desde finales del siglo XX: la conservación y el uso sostenible de la biodiversidad, campo en el que el Instituto Franklin de la UAH viene trabajando con análisis comparados transatlánticos desde sus orígenes, y muy especialmente desde 2001, a través del programa «Friends of Thoreau».

2. ESTRUCTURA DE LA INVESTIGACIÓN

Esta investigación, como se ha anunciado antes, se ha hecho realizando un análisis comparado de dos enfoques distintos: el que existe en la Unión Europea y el de Estados Unidos, y se hace hincapié, cuando procede, en las diferencias o particularidades de cada uno.

Teniendo en cuenta esta doble perspectiva, este trabajo parte de una previa contextualización, es decir, de una explicación de: (a) lo que ha supuesto en general la digitalización respecto de la economía del conocimiento y, más en concreto, lo que ha supuesto la digitalización de datos científicos en la medida en que ha dado lugar a la creación de nuevas infraestructuras y a cambios necesarios en los derechos existentes para la protección de la «propiedad» de los datos científicos de general; y (b) de lo que se entiende por políticas de biodiversidad.

2.1. Dos marcos de contextualización

A. La revolución digital

El primero de los marcos contextuales aborda la aparición de las nuevas tecnologías y el impacto que esto ha tenido a todos los niveles.

A.1.- La aparición de un nuevo paradigma: Internet

El análisis de este marco contextual arranca, como no podía ser de otra manera, en los primeros pasos de lo que hoy es Internet, es decir, un proyecto iniciado como una forma de garantizar las comunicaciones en plena Guerra Fría, que fue evolucionando, primero, en una herramienta para facilitar datos e investigaciones subvencionada por la National Science Foundation de Estados Unidos hasta llegar al momento en que la red de redes se configuró con la estructura que hoy tiene.

No cabe duda de que Internet representa un cambio en la gestión y generación del conocimiento. Ha supuesto un cambio de paradigma. Al igual que en su día se habló de la revolución industrial, hoy se habla de la revolución digital o de «Tercera Revolución Industrial» (Rifkin, 2012).

De hecho, la era de Internet se conoce también como la era del conocimiento, la sociedad del conocimiento, la economía del conocimiento, etc., lo que prueba que este cambio de paradigma ha permeado la economía productiva a todos los niveles.

¿Qué significan estos términos? ¿Cuáles son las implicaciones, las características de esta nueva era o revolución? La velocidad, la extensión geográfica y el volumen de generación de conocimiento mediante la transposición de los datos y su crecimiento exponencial hacen que, de las primeras compilaciones novedosas, aunque estáticas, de *data warehouse*, se haya pasado a lo que hoy se conoce como *big data*. Sin duda, un cambio que hace necesario reformular las herramientas existentes hasta este momento.

A.2.- La gestión del conocimiento

Se plantean, así, toda una serie de cuestiones. Tal vez, la primera de todas ellas sea la de definir la forma en que se gestiona este conocimiento. De repente, las formas de gestión del conocimiento parecen haber quedado obsoletas debido a los volúmenes de datos y de sus expresiones (escritas, gráficas, audiovisuales, etc.) que se gestionan hoy, así como a la manera a través de la cual es necesario acceder a los mismos.

Ello ocurre tanto para el acceso y reutilización de los datos sociales necesarios para la racionalización de las políticas públicas como para los datos científicos (surge así el concepto de e-ciencia); y para mejorar los unos y los otros aparecen las infraestructuras a las que se aplican distintos apellidos: infraestructuras intelectuales, infraestructuras del conocimiento, e-infraestructuras, etc.

Respecto al concepto de «dato científico», se trata de un concepto acuñado *ad hoc* para esta tesis, aunque también lo usan las instituciones españolas y latinoamericanas con asiduidad. Podría más bien hablarse de «datos de los procesos de investigación», por traducir más literalmente la reinante expresión de la lengua inglesa (*research data*), que sí es un concepto perfectamente generalizado en la literatura escrita en inglés). En cualquier caso, debe quedar claro que abarca tanto los datos de las ciencias duras (y menos duras), las ciencias sociales y las humanidades como los más difíciles de hacer interoperables: los de las ciencias inter y multidisciplinares (Alonso García, 2016b). La clave no está tanto en la denominación en sí, sino más bien en su construcción conceptual. Se hace necesario analizar sus orígenes y las funciones que cumplen.

Aunque la construcción conceptual es parecida, existen dos sistemas jurídicos que enfocan la definición de los conceptos de forma muy distinta. Por un lado, está la construcción europea (donde se habla más bien de e-infraestructuras) y, por otro, la estadounidense (que habla, más comúnmente, de ciberinfraestructuras). La función principal, en ambos casos, es la gestión de los datos y en particular de los datos científicos a efectos de su utilización por parte de las infraestructuras de e-ciencia. La forma de abordarlos en Estados Unidos y en Europa es, sin embargo, diferente, más allá del nombre.

A.3. Regulación de las e-infraestructuras: diferencias entre la Unión europea y Estados Unidos

En Europa se crean unas estructuras concretas, con una detallada regulación jurídica para abordar estas cuestiones. Aparece el Foro Estratégico Europeo sobre Infraestructuras de Investigación (ESFRI, por las siglas en inglés de European Strategy Forum on Research Infrastructures), que sirve de marco para la creación de la estrategia europea en lo que respecta a las infraestructuras de investigación. Como un paso más, se crean los Consorcios de Infraestructuras de Investigación Europeas (ERIC, por las siglas en inglés de European Research Infrastructure Consortium), que son un nuevo tipo de persona jurídica, no sujeta al derecho nacional, sino al derecho internacional y supranacional europeo, para cuya constitución se requiere el acuerdo como mínimo de tres países —en origen, los tres tenían que ser miembros de la UE, pero en la actualidad esto no es necesario—.

Los ERIC son una muestra de la forma continental de regular en Europa. Se adoptan una serie de normas concretas que regulan qué proceso deben seguir para constituirse, de qué manera tienen que funcionar y qué condiciones tienen que cumplir. Se les dota de personalidad jurídica para que puedan actuar en el espacio internacional de forma autónoma. Se simplifica y ordena así la creación de infraestructuras de investigación en el marco de la Unión Europea. Lo más importante es que esta regulación europea —el Reglamento [CE] 723/2009— (Consejo de la Unión Europea, 2009) evita tener que negociar, celebrar y ratificar un tratado internacional *ad hoc* por cada una de las partes contratantes (Estados) que formen parte de la infraestructura de investigación o consorcio, que es lo que era usual hasta entonces. Véanse, por ejemplo: *el European Molecular Biology Laboratory* (EMBL), *el Conseil Européen pour la Recherche Nucléaire* (CERN), *el International Thermonuclear Experimental Reactor* (ITER) o *la Global Biodiversity Informatics Facility* (GBIF), entre otros muchos, todos los cuales se rigen por su correspondiente tratado internacional.

La construcción de estos ERICS obedece a la necesidad de armonizar sistemas jurídicos y culturas diferentes mediante instituciones supranacionales. Su

«particularidad», en realidad, es que son una consecuencia lógica de la propia naturaleza —supranacional— de la Unión Europea y del hecho de que las economías de escala son mucho más ventajosas a ese nivel que al nacional, donde muchos Estados miembros no podrían tener estas e-infraestructuras.

El sistema de Estados Unidos es completamente distinto al europeo. No existen marcos de referencia generales, sino construcciones *ad hoc*, lo que nos ha obligado a analizar en detalle en qué consisten estas ciberinfraestructuras, de qué manera se articulan y, sobre todo, profundizar en alguna de las organizaciones que promueven la investigación mediante la concesión de *grants* o subvenciones, como la National Science Foundation o el National Institute of Health, lo que implica un análisis de los límites al poder federal en la utilización de lo que se conoce en derecho constitucional como el *granting power*.

A.4. Las cuestiones jurídicas que plantea la revolución digital

Además, la revolución digital no solo ha dado lugar a nuevas instituciones, sino que su operatividad misma (el acceso a los datos digitales y su flujo), también ha necesitado cambios jurídicos sustanciales.

Es fundamental entender el nuevo marco y estudiar cuáles son los problemas jurídicos que se plantean en la interoperabilidad de datos, es decir, lo que se conoce con el nombre de interoperabilidad legal que antes se ha descrito.

¿Cuáles son los problemas jurídicos de partida que plantea la tecnología? De alguna manera, el dato es la base de este sistema, pero ¿cuál es el concepto de dato? ¿Qué son los datos de investigación?

Por consiguiente, el primero de los problemas jurídicos que se plantean en el funcionamiento de las e-infraestructuras (o ciberinfraestructuras) es el de la interoperabilidad legal; ello obliga a definir el concepto de dato y las particularidades que presenta la mera reproducción de la realidad a través de la tecnología digital, en especial cuando a través de estas ciberinfraestructuras se multiplica el uso de dichos datos hasta extremos ni siquiera imaginables hace unos años.

Entendido, si es que se logra, cuál es el estatuto jurídico de los datos digitales, el segundo de los problemas consiste en analizar hasta qué punto los obstáculos jurídicos pueden dificultar o hacer prácticamente imposible su flujo (*workflow*). A su vez, no se puede entrar a fondo en el análisis de qué es la interoperabilidad y cómo se logra la legal sin estudiar antes el objeto de esa interoperabilidad: el dato científico digital.

Por ello, dicho objetivo se pospone a secciones posteriores para adentrarnos previamente en todos los problemas que plantea la delimitación de qué es un dato

digital, qué es un dato digital científico y cuál es su régimen jurídico. ¿Puede protegerse mediante técnicas de derechos de propiedad intelectual o es un «dato abierto», es decir, común de la humanidad y no protegible?

¿Cuáles son las cuestiones que se plantean en lo que respecta a la propiedad y a la protección de los datos? ¿Son los datos susceptibles de apropiación? ¿Qué herramientas pueden utilizarse para proteger los datos? ¿Deben abandonarse totalmente, en la moderna economía del conocimiento, unos instrumentos como son los genéricos derechos de propiedad intelectual o *copyright*? Estos, si bien sirvieron como instrumento esencial para la primera y segunda revolución industrial, ¿frenan la economía de la tercera revolución, de la revolución del conocimiento?

Así pues, no solo es relevante la definición de qué es el dato o qué son los datos de investigación cuando se digitalizaban; hay que ir más al detalle, ha sido necesario analizar la cuestión de la propiedad de los datos y las formas de protegerlos, y es ahí donde vuelve a surgir el desencuentro: hay que tener en cuenta que, en principio, aunque haya habido tratados internacionales, Europa y Estados Unidos tienen una regulación muy distinta de dichos derechos de propiedad intelectual o *copyright*.

Esta diferencia no se da solo con respecto a los derechos de propiedad intelectual o *copyright* sobre los propios datos, sino también sobre las técnicas de organización de los mismos, en concreto, los derechos de propiedad intelectual o *copyright* sobre las bases de datos, cuyas regulaciones difieren de forma radical en ambos ordenamientos jurídicos.

Dentro del modelo europeo de protección de los datos (de investigación) existen tres niveles de protección: el primero es la protección de los datos en sí mismos, es decir, del contenido (por ejemplo: una fotografía de un ejemplar de una especie amenazada, ¿puede estar sujeta a *copyright*?).

El segundo nivel es el de la protección o falta de ella mediante derechos de propiedad intelectual de la manera de estructurar los datos (es decir, la protección de las bases de datos en sí mismas) como una protección de los derechos de autor para la creación intelectual constituida por la selección o disposición de materias.

En tercer lugar, una protección *sui generis* que garantiza la protección de una inversión (financiera, en recursos humanos, esfuerzo y energía) efectuada en la obtención, la verificación o la presentación del contenido de una base de datos que ocurre en Europa a través de lo que se conoce como los derechos *sui generis* sobre bases de datos, un derecho de la normativa europea sin parangón en la normativa estadounidense.

Por último, se analiza el proyecto de directiva que crea un nuevo derecho de propiedad intelectual, la licencia para llevar a cabo el llamado *text and data mining* (en

adelante, TDM) y las implicaciones que la misma puede tener en lo que respecta a la protección de los datos y, sobre todo, de las barreras presuntamente irracionales que estas nuevas licencias TDM van a suponer para el flujo normal de datos a gran escala. Es una norma, aún en construcción, que puede tener un alto impacto en la regulación actual del flujo de los datos. ¿Qué problemas plantea y desde qué perspectiva lo aborda?

En lo que respecta al modelo estadounidense, es importante hacer el análisis de la evolución: ¿cómo se configura la protección de las bases de datos en Estados Unidos? ¿Existe la protección del continente (bases de datos) como ocurre en el caso de Europa? O, ¿solamente protege el contenido, la propiedad del derecho? ¿Qué nivel de originalidad es exigible? ¿Cómo se aplica el *fair use* como excepción a los excesos de susceptibilidad de apropiación de los datos?

Además, es necesario analizar cuál de los dos sistemas es más adecuado, el de la Unión Europea o el de Estados Unidos, y si es necesario un sistema más global para el progreso de la ciencia y, en especial, para el de la ciencia aplicada mediante técnicas digitales y, por tanto, si son esencialmente compatibles ambos sistemas.

Volviendo a los derechos de propiedad intelectual o *copyright*, estos nacieron como una forma de garantizar la innovación, las obras intelectuales y artísticas, pero no fueron pensados para la publicación de los datos. Puede que hoy se hayan convertido en una losa frente al fenómeno de la digitalización —el primero de los marcos contextuales que se tratan en esta investigación— al impedir dicha innovación o hacerla prácticamente imposible.

¿Son útiles hoy en día para la organización y transmisión de datos digitales? Piénsese que, por ejemplo, los datos pueden estar en una publicación científica digitalizada. Es clave, en este momento, el «acceso en abierto» (*open access*). Existe un fuerte movimiento social y político que lo defiende. ¿Qué es el acceso abierto? ¿Cómo encaja con los derechos de propiedad industrial e intelectual?

Con esta preocupación de una necesidad de soluciones más globales, empiezan a aparecer nuevas herramientas o nuevos enfoques: los conocidos con el nombre de *open* (abiertos): el *open access*, el *open data*, la *open science*, etc., que en Estados Unidos se engloban bajo la denominación de *public access*.

¿Qué es el *open/public access* o acceso en abierto? ¿Qué fases han existido en la creación de las políticas de acceso abierto? ¿Cuál es la definición de este concepto?

El consenso en lo que respecta a los denominados FAIR Principles, es una cuestión discutible (Wilkinson *et al.*, 2016): se plantea la cuestión de si los datos deben ser *findable*, *accessible*, *interoperable* y *reusable*.

Respecto a la forma, también digital, de plasmación de los datos, el problema se traslada a dichas formas de expresión, como ocurre con el ejemplo de datos en revistas que se acaba de señalar y, sobre todo, al *software* que permite acceder y reutilizar los datos, lo que se debe a que muchas veces la única forma posible de acceder a los mismos o de reutilizarlos consiste en utilizar, a su vez, el *software* concreto y específico en el cual los mismos están recogidos. Por tanto, aunque los datos no estén sujetos a licencias, el instrumento de acceso a los mismos sí lo está, de manera que, por ejemplo, los datos pueden ser *open* pero el sistema de acceso a los mismos estar cerrado (*closed*), según los países, está protegido por *copyright*, por patentes. Por eso surge el *open source*. ¿Qué es el *open source*? ¿Qué significa usar *open source* referido al *software*? ¿Cuáles son las implicaciones que puede tener el que sea *open* frente a que esté sobre dicha licencia?

De nuevo, en lo que respecta a estos puntos, existen dos modelos o sistemas a la hora de abordar estas cuestiones: el europeo y el estadounidense. Se hará un análisis más detallado de los que ambos plantean.

¿Pueden considerarse el *open/public access* y el *open source* la base de la construcción de la interoperabilidad legal global?

Parece, sin embargo, que existen dos fuerzas contrarias que chocan: si bien una parte de la doctrina y de la regulación defiende la aplicación de estos principios con la finalidad de acelerar el desarrollo humano, otra parte de la doctrina se aferra a los viejos paradigmas más restrictivos defendiendo la ocultación de datos científicos como forma legítima para el avance de la ciencia al proporcionar a los científicos (y a las instituciones que se mueven en torno a ellos) incentivos en forma de rendimientos económicos y carreras profesionales prestigiosas que solo se obtienen reforzando los derechos de propiedad intelectual o *copyright*. Esto se refleja incluso en normas, como puede ser el caso de la propuesta de directiva de *text and data mining* europea, todavía en discusión, y a la que ya se ha hecho referencia un poco antes, que, en lugar de abrir la ciencia, la cierran todavía más con una nueva licencia que es necesario obtener.

¿De qué manera se aborda en la Unión Europea? Se analizarán, en este punto, el marco legislativo y las iniciativas políticas de *open access* y sus contradicciones internas con las licencias TDM. Además, se estudiará la Agenda Digital para Europa con la finalidad de ver hacia dónde se dirige, así como la creación y el funcionamiento del mercado único digital.

El modelo estadounidense parte de un concepto muy diferente, de las llamadas políticas de *public access*, una creación de la Administración Obama. ¿En qué consisten las políticas de *public access*? ¿Cuál es el origen de las mismas? ¿Cómo se han utilizado? En el proceso, se analizarán las distintas normas de aplicación.

A.5. Data sharing

Después del análisis de ambos modelos o sistemas de abordar estas cuestiones en abstracto, la presente tesis procede a analizar todas las cuestiones de intercambio de datos o *data sharing* en dos instituciones globales: GEOSS y CODATA, instituciones ambas enfocadas principalmente a compartir datos de investigación o científicos.

GEOSS

Respecto a GEOSS, el instrumento de gestión de datos, de imágenes satelitales o Sistema de Observación de la Tierra (GEOSS, por las siglas en inglés de Global Earth Observation System of Systems), nace como un grupo creado *ad hoc*, intergubernamental, copresidido por la Comisión Europea, Japón, Sudáfrica y Estados Unidos. Su misión fue la de elaborar un plan a diez años (presentación en febrero de 2005) y su objetivo principal era la coordinación en materia de observación de la tierra.

El sistema de sistemas tiene una función integradora que persigue dar coherencia a todas las recopilaciones de datos existentes en este momento. Su finalidad principal es la interconexión de sistemas.

Se define como una forma de dar coherencia a la observación y de registrar valores. Uno de los posibles efectos será evitar, a futuro, que fenómenos «observados» tengan efectos devastadores para el hombre por no existir los canales de información adecuados.

CODATA

Respecto a CODATA, un comité científico interdisciplinar, Consejo Internacional de Ciencia (conocido como ICSU por sus siglas en inglés), trabaja para mejorar la calidad, la confiabilidad, la gestión y la accesibilidad de los datos importantes en todos los ámbitos de la ciencia y de la tecnología.

Identifica tres áreas de prioridad. La primera es la promoción de los principios, políticas y prácticas necesarias para la construcción del *Open data* y del *Open science*. La segunda es el avance de las fronteras de los datos científicos. La tercera, la construcción de la capacidad necesaria para la *Open science* mediante la mejora de la gestión de los datos y las funciones de los sistemas de ciencia nacionales necesarios para sostener el *Open data*.

¿Es la colaboración, compartir los datos en abierto, como muestran las dos instituciones (GEOSS y CODATA), el nuevo paradigma de hoy, lo que permitirá un avance de la ciencia mayor que el desarrollado hasta ahora?

¿Están frenando los regímenes actuales de derechos de propiedad intelectual o *copyright* las necesidades de los científicos al limitar la interoperabilidad de los datos? Es decir, ¿frenan la posibilidad de que una máquina pueda utilizar datos que están en otra máquina (*machine-machine readable data*)?

La interoperabilidad legal

La interoperabilidad legal es un concepto novedoso. La clave consiste en dar acceso, idealmente hasta su reproducibilidad total, a los datos desde el punto de vista jurídico, de manera que —también otro ideal— las máquinas mismas puedan juzgar las limitaciones jurídicas que existen al flujo de datos sin tener que consultar a un abogado o remitir necesariamente a quien quiere reutilizarlos a consultar, bien a quien ha documentado la cesión de derechos de propiedad intelectual o *copyright*, bien a su propio abogado, bien a quienes han firmado el contrato o documento a los que la máquina les remita.

Es más, incluso hay nuevas formas de contratación que pueden consistir en hacer clic con el cursor sobre alguna de las opciones que remiten a condiciones estandarizadas de contratación que la propia máquina debe saber interpretar.

Alternativamente, cuando ello sea imposible, la máquina debe saber cómo puede identificar —en el resultado final de los nuevos datos surgidos del uso de los anteriores mediante nuevas aplicaciones o productos creados por las infraestructuras— qué datos se han perdido, cuáles no son enteramente fiables o, simplemente, si no se tienen en cuenta debido a que las condiciones que las hacían originariamente no utilizables impidieron a las máquinas enjuiciar si dichos productos o aplicaciones podrían en hipótesis tener límites para su utilización de datos o bases de datos jurídicas.

La construcción de la interoperabilidad legal y de los cambios de paradigma en los que este mundo está inmerso requiere de una aproximación diferente. Este concepto de interoperabilidad legal es más amplio y va más allá, por tanto, del clásico concepto de interoperabilidad legal, que es más limitado y que se basa en hacer llegar al usuario del dato las limitaciones que puede tener su acceso, descarga, utilización o reutilización. Es decir, la interoperabilidad legal en un principio se estructuró como un sistema a través del cual en la web misma se localizan las condiciones jurídicas vinculadas a los datos. Actualmente, esto resulta totalmente inútil, puesto que cuando se pasa del manejo de datos o bases de datos al manejo de *big data*, se hace imposible este último si la interoperabilidad legal no va más allá de la identificación de los problemas jurídicos de la utilización de los datos. Esto es así porque el sistema se colapsaría si hubiera que reproducir esas condiciones en miles o millones de datos, con su especificidad jurídica para cada uno de esos datos o bases de datos.

Para contribuir a la simplificación y mejor comprensión de la necesidad de interoperabilidad legal ha nacido el movimiento *copyleft*, como una forma de permitir a los autores, no especializados en cuestiones técnicas de propiedad intelectual, poner los trabajos y datos de su propiedad a disposición de todos para su reutilización. Un ejemplo que no podía dejar de mencionarse en este trabajo son las licencias Creative Commons.

Además, distintas organizaciones de distintos ámbitos (nacionales, regionales, globales) empiezan a analizar las cuestiones que surgen en la aplicación de estos derechos de propiedad intelectual y a tratar de contribuir a su resolución mediante la publicación de principios, reglas y recomendaciones tendentes a mejorar la publicación de los datos y la reutilización de los mismos. Se analizarán también, en detalle, algunos de estos principios y, en concreto, los de la Research Data Alliance (RDA) y sus *Implementation Guidelines*, principios de interoperabilidad legal publicados en octubre de 2016 en el marco de la RDA en cuya elaboración ha participado la autora de este trabajo.

También se analizarán, como un paso más, los Principios del Foro de Belmont y, los Principios Panton (*Principles for Open data in Science*) que parten de la asunción de que la ciencia está basada en la construcción, la reutilización y la crítica abierta de las publicaciones científicas. Finalmente, se analizarán los llamados *FAIR Guiding Data Principles* (*findable, accessible, interoperable, and re-usable*).

Por último, se analizará el ámbito del dominio público (*public domain*) y lo que ello supone como alternativa al *copyleft* o a las licencias Creative Commons. ¿Podría ser esencial como paso previo a la construcción de un concepto de interoperabilidad legal global? Una parte de la doctrina, de hecho, reclama que sea necesario publicar los datos científicos sin restricciones e incluso ponerlos o «dedicarlos a» el dominio público como alternativa a estas licencias *copyleft*. ¿Cuál es el interés general que debe predominar y cómo se protege mejor?

¿Puede hablarse por lo tanto de interoperabilidad global? ¿Es una realidad o una utopía? ¿Deben considerarse estos proyectos como meros precursores de una visión más global?

B. La biodiversidad

El segundo marco de contextualización de este trabajo es el mundo de la biodiversidad y su conservación y uso sostenible.

Este punto arranca de la consolidación de la construcción histórica del concepto de biodiversidad en la última década del siglo XX. Su elaboración, iniciada a partir de los años setenta, parte de la toma de conciencia por parte de la humanidad de la necesidad de protegerla conservándola, sin perjuicio de hacer de ella un uso

sostenible frente a su pérdida acelerada, consecuencia de la huella que los comportamientos del hombre están produciendo sobre la misma (la sexta extinción masiva de especies, véase la figura que sigue, donde están las cinco primeras) y ante la evidencia, hoy compartida, de que el ser humano y la biodiversidad están indefectiblemente unidos. Su supervivencia misma está unida a la protección de los distintos ecosistemas y de los servicios de los mismos.

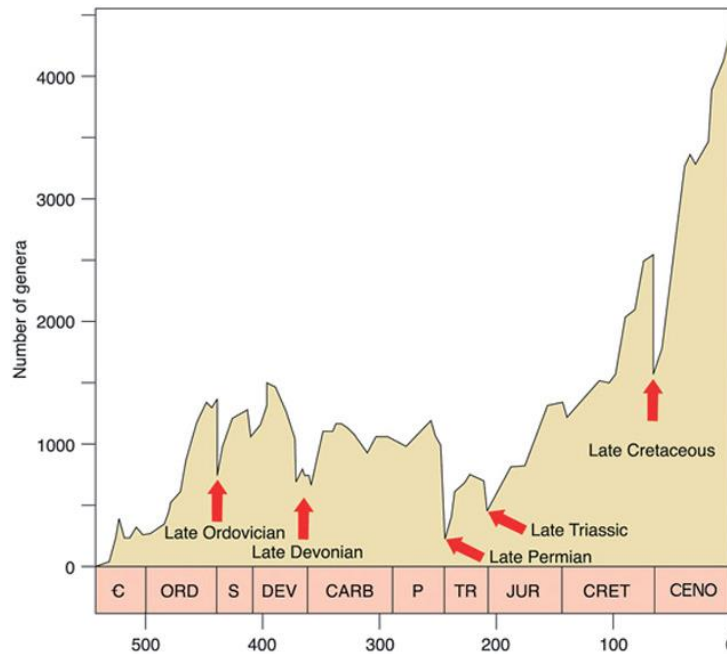


Figura 1. Tiempo geológico en millones de años

Aunque hay múltiples definiciones de «biodiversidad» (término abreviado de «diversidad biológica»), se utiliza en esta tesis doctoral la más común, la del artículo 2 del Convenio de Naciones Unidas de Diversidad Biológica (en adelante, CDB), si bien ampliada con el añadido de la protección de los paisajes ecológicos, ámbito de acción un poco más amplio que el ecosistema.

Dicho artículo 2 establece que: «Por “diversidad biológica” se entiende la variabilidad de organismos vivos de cualquier fuente, incluidos, entre otras cosas, los ecosistemas terrestres y marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte; comprende la diversidad dentro de cada especie, entre las especies y de los ecosistemas». Hoy en día el conjunto de ecosistemas agrupados en un paisaje ecológico o incluso en una biorregión puede variar, por lo que la diversidad de paisajes ha pasado a engrosar la lista, ya que los mismos ecosistemas en un ambiente seco ofrecen una riqueza biológica muy distinta si aumenta la humedad, por ejemplo.

Explicar en detalle cuáles son y en qué consisten las políticas y el derecho de la conservación y uso sostenible de la diversidad biológica iría mucho más allá, en sí mismo, de una tesis doctoral. Bastaría por tanto remitir, lo que se hace ya por

adelantado, a los capítulos 3 —«The Biodiversity convention. Access to genetic resources and benefitsharing (ABS). The sister conventions (endangered and migratory species —CITES, CMS—, desertification —UNCCD—)»— y 4 —«Protecting Wetlands: The Ramsar Convention»— del libro de Alonso García, *Introduction to International Environmental Law: Handbook with Cases and Materiales for American Lawyers* (2008)¹ y a la página web del CDB y los restantes cuatro tratados internacionales que integran el «liaison group» internacional de protección de la biodiversidad: el Convenio de Ramsar (UNESCO, 1971), CITES (Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres, 1973), Patrimonio Mundial de la UNESCO e IPGRI and the FAO y el International Treaty on Plant Genetic Resources for Food and Agriculture (FAO, 2009).

En lo que sí se ha adentrado más en presente trabajo de investigación es en el entramado institucional que intenta poner al servicio del CDB la ciencia en la que sus órganos decisorios deberían basar sus decisiones, sin perjuicio de que en este campo, como parte del derecho ambiental, se aplica el principio de precaución y, por tanto, la incertidumbre científica no debe bloquear la toma de decisiones, sino, por el contrario, legitimarla.

A diferencia de la Convención Marco de las Naciones Unidas sobre el Cambio Climático (UNFCCC, por las siglas en inglés de United Nations Framework Convention on Climate Change) también del año 1992, en la que este problema fue solucionado desde su origen, y de raíz, mediante la creación del Panel de Cambio Climático (IPCC, por las siglas en inglés de Intergovernmental Panel on Climate Change), la gestión de la interrelación entre la ciencia y la toma de decisiones en materia de biodiversidad a nivel global ha carecido de modelo.

La gran diferencia entre el sistema de lucha contra el cambio climático y el de la lucha por la conservación y el uso sostenible de la biodiversidad es que el marco científico en el cual se produce el análisis de los problemas relacionados con el cambio climático es mucho más lineal y, por tanto, más sencillo de abordar que el de la pérdida de biodiversidad.

El primero, es decir, el de cambio climático es quizás más complejo en cuanto a la negociación de soluciones para abordar los problemas derivados del análisis científico, ya que está en juego el modelo económico productivo en su casi totalidad. Sin embargo, en el segundo, el de la biodiversidad, el problema radica en la práctica inexistencia de sistemas científicamente consensuados para analizar la pérdida de biodiversidad misma (si se produce o no y a qué escala espaciotemporal) y sus

¹ *Friends of Thoreau Program Series*, 3.ª ed, núm.1, Madrid: URJC-Friends of Thoreau, 2011.

implicaciones: es un problema mucho más complejo dado que debe abordarse analizando datos desde el nivel genético hasta el nivel espacial del paisaje ecológico o incluso del conjunto del planeta, y tienen que examinarse conjuntamente con datos que sirven de indicadores físicos y químicos así como socioeconómicos, que inciden sobre la biodiversidad y tienen mucha mayor indefinición y extensión.

El CDB

No obstante, la tesis resume la gestación del CDB y el posicionamiento de Estados Unidos frente al mismo, que merece una atención especial, analizando de forma separada las razones por las que Estados Unidos decidió (posición que mantiene) no ratificar el CDB. En efecto, Estados Unidos, aunque implicado en la creación o el desarrollo teórico del mismo, no acaba de ratificarlo.

Esto es así porque lo que sí ha resultado necesario es analizar su estructura: ¿cuáles son los órganos que integran el «régimen» internacional que se ocupa de que el CDB funcione? Cada uno de ellos merece una consideración separada en lo que respecta a su constitución y la forma en la que desempeña su rol en el marco del convenio.

¿Cuál es el objetivo del CDB? ¿Ha cumplido las expectativas con las que se creó? ¿Cuáles son sus principales retos estratégicos? ¿Qué son los normalmente conocidos como *Aichi targets*?, entre otras cuestiones, son preguntas necesarias para identificar los objetivos que debe perseguir la revolución digital en este campo concreto de la conservación y el uso sostenible de la biodiversidad.

Más en concreto, lo que ha resultado más importante para la tesis es contestar a la siguiente pregunta: ¿cuál es el entramado institucional puesto en marcha por el CDB para abordar las dudas científicas que permiten abordar, a su vez, todos estos problemas institucionales? Basta con examinar el complejo cuadro de instituciones involucradas (órganos del «régimen» del CDB y organizaciones externas al mismo) en dar solución a nivel global a la pérdida de biodiversidad y en buscar una solución mínimamente adecuada a la misma (por ejemplo: el principio de *No net biodiversity loss*) para llegar a la conclusión de que el problema es bastante más complejo que el que se intenta abordar en la Convención de las Naciones Unidas en la Lucha contra el Cambio Climático (UNFCCC).

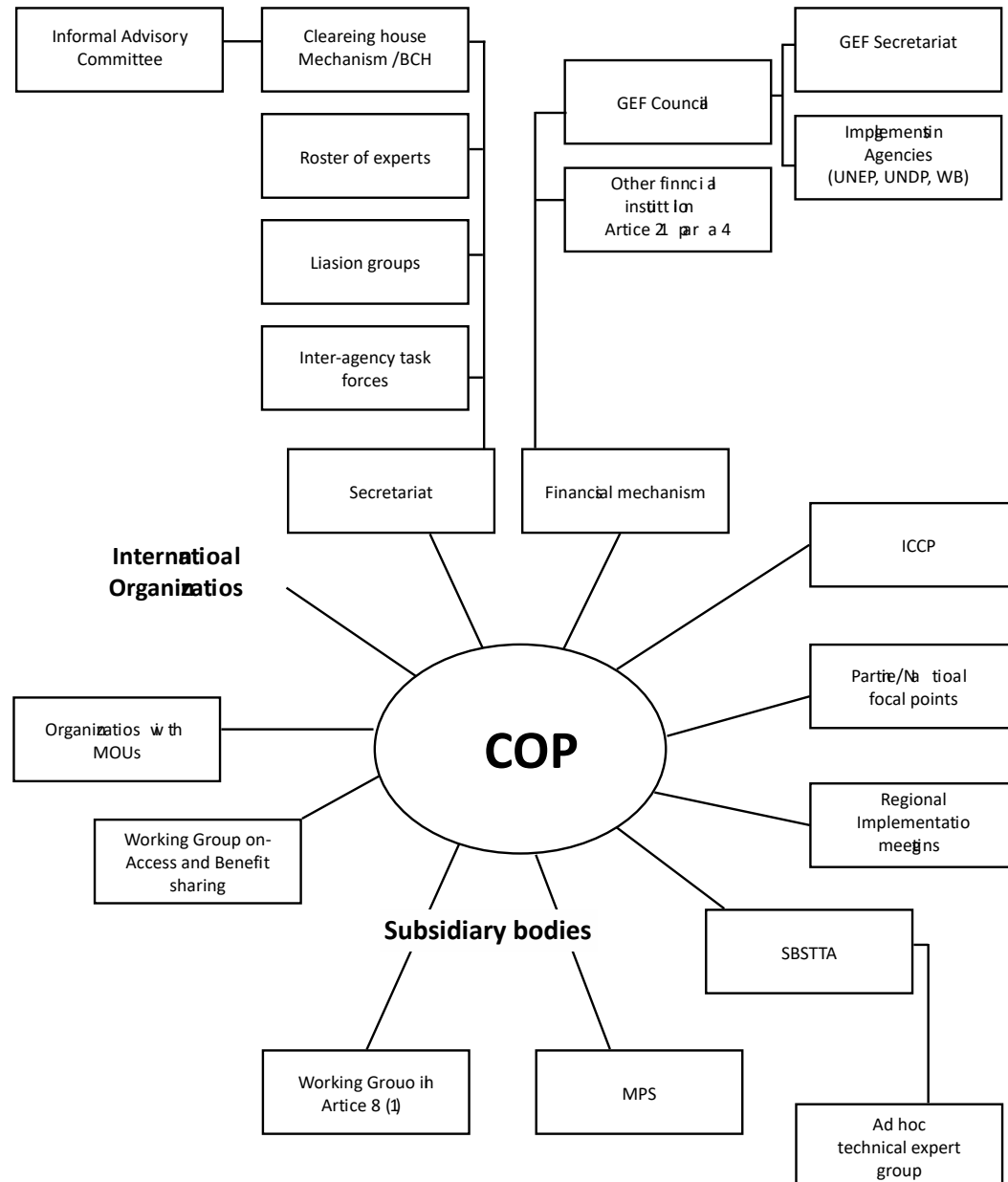


Figura 2. Cuadro sintético de los órganos del CDB. Fuente: Alonso García, 2008.

¿Resulta suficiente ese entramado institucional para la consecución de los objetivos con los que nace?

Una especial consideración ha merecido el estudio del Subsidiary Body on Scientific, Technological and Technical Advice (SBSTTA). En efecto, si bien la creación de este órgano en el marco del CDB tenía como objetivo alinear la ciencia y la política, no ha sido suficiente. A pesar de la creación del SBSTTA, el CDB no haya, logrado que sea el marco científico necesario en el que realizar aportaciones relevantes y consensuadas y, consecuencia de ello, se ha dificultado la tarea de identificar objetivos para frenar la pérdida de biodiversidad ¿Cuáles serían las causas de no

haberlo alcanzado? ¿Exceso de ambición, inadecuada implementación o falta de consenso acerca de cómo medir dicha pérdida?

Realmente, no existe un órgano equivalente al IPCC del régimen de cambio climático para la biodiversidad y ello es clave para que el proceso del SBSTTA no haya llegado a cumplir para la biodiversidad lo que el IPCC ha logrado para el cambio climático.

Como consecuencia de las insuficiencias del SBSTTA para abordar los retos científicos que supone la pérdida de biodiversidad, un nuevo instrumento ha intentado recoger el testigo: la Plataforma Intergubernamental sobre Biodiversidad y Servicios de los Ecosistemas (IPBES, por las siglas en inglés de Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services), persona jurídica internacional creada bajo el auspicio de las Naciones Unidas, en una reunión en Panamá en 2012 por más de cien Estados.

El IPBES

El IPBES, al igual que el CDB y su SBSTTA, pretende replicar el Panel Internacional de Cambio Climático (IPCC), sin lugar a dudas, la herramienta que mejor ha funcionado y que más compromiso ha alcanzado como instrumento internacional en la interrelación entre ciencia y política.

El IPBES debe afrontar una serie de retos, con la finalidad de lograr lo que el CDB no consiguió y como apoyo al mismo. Como plataforma que es, el IPBES demuestra mucha más flexibilidad a la hora de abordar las cuestiones científicas relacionadas con la biodiversidad. Además, su configuración permite trabajar en colaboración con muchos otros instrumentos. De hecho, es cuestión de tiempo que el IPBES sea alimentado por todos los datos de las distintas herramientas analizadas en el punto siguiente.

2.2. La aplicación de las nuevas tecnologías surgidas de la revolución digital a la ciencia de la biodiversidad: la bioinformática

Con esta reflexión y después de haber llevado a cabo una doble contextualización (nuevas tecnologías y biodiversidad), se inicia la segunda parte de la presente tesis analizando en profundidad el impacto de la aplicación de las nuevas tecnologías a la conservación de la biodiversidad. Se construye así el concepto de informática de la biodiversidad o bioinformática. Se dedica una primera reflexión específica a los orígenes y construcción de esta disciplina.

La bioinformática

¿Qué es la bioinformática? ¿Cómo se está construyendo? ¿Es una nueva disciplina o una disciplina integradora de distintas ramas? Esta materia novedosa analiza la forma de aplicar todos los nuevos paradigmas de gestión de datos científicos y de investigación a los datos de biodiversidad.

Como punto de arranque ha sido necesario analizar el concepto de «bioinformática» que fue creado a efectos de conseguir la aplicación de las tecnologías digitales y la utilización masiva de los datos digitalizados sobre biodiversidad, a escala global, para, por lo menos, tener conocimiento del estado en que se encuentra la misma; es decir, cuáles y de qué tipo tienen que ser los datos digitalizados relevantes de la biodiversidad misma existente en nuestro planeta desde el nivel genético al del paisaje ecológico (pasando, como ya se apuntaba, por la biodiversidad de individuos y poblaciones de una misma especie, la diversidad de especies, la diversidad de ecosistemas, la diversidad de hábitats y, finalmente, los paisajes ecológicos).

La creación del GBIF

Clave en este proceso fue la creación por parte de un conjunto de países y organizaciones supranacionales, entre los que se encontraban Estados Unidos y la Unión Europea, de la denominada Global Biodiversity Information Facility (GBIF), una organización internacional creada por un Memorando de Entendimiento internacional entre varios Estados que daba cumplimiento a una recomendación de la OCDE de 1999. El GBIF se configura como una de las primeras herramientas en la construcción de la bioinformática.

El GBIF es una organización intergubernamental estructurada en nodos. Su objetivo es facilitar el acceso —vía Internet, de manera libre y gratuita— a los datos de biodiversidad de todo el mundo para apoyar la investigación científica, fomentar la conservación biológica y favorecer el desarrollo sostenible. Nació con el objetivo de aplicar la informática como mecanismo para facilitar y administrar información proveniente de la naturaleza. ¿Cómo colabora el GBIF con el IPBES?

Sin embargo, antes de analizar dicha relación con el IPBES, ha sido necesario estudiar cuál es el *statu quo* concreto en que a fecha de hoy se encuentra la bioinformática y hasta qué punto el GBIF, como principal organización internacional que pretende crear instrumentos aplicados a la conservación de la biodiversidad a través de la misma, tiene carencias o defectos estructurales o coyunturales que le impiden alcanzar esos fines. Esto, además, parece ocurrirle tanto al GBIF como a otras infraestructuras.

Tras muchos años de funcionamiento, el GBIF, con el impulso del proyecto CReATIVE-B al que se hará referencia más abajo, procedió recientemente a evaluar

sus logros dentro del marco más amplio de los progresos que la bioinformática había tenido hasta 2012. Para ello organizó la Global Biodiversity Informatics Conference (GBIC),² donde se reunió a cien expertos en biodiversidad informática en Copenhague para analizar tanto las ineficiencias del propio GBIF como las de GEO BON y otros sistemas (LifeWatch, ALA, SANBI, etc., a los que se aludirá más adelante).

Dicha conferencia reunió a científicos de todos los ámbitos: biólogos, geólogos, informáticos, etc., e incluso juristas. A resultas de la misma, se publicó en el año 2013 el *Global Biodiversity Informatics Outlook (GBIO): Delivering Biodiversity knowledge in the information age*.

¿Cuáles son las conclusiones del GBIO y los planteamientos que hace a futuro? Por primera vez, se realizó un análisis del punto de partida y de las áreas sobre las que es más necesario profundizar. Se han recogido en el presente trabajo las conclusiones del GBIO y el punto de desarrollo en que se encuentra cada una de las distintas áreas.

Por consiguiente, si bien el IPBES se constituye en la organización internacional de apoyo científico por excelencia al CDB, ya existe para el campo de la bioinformática el GBIF como organización internacional más especializada en bioinformática; institución que, por lo demás, es plenamente consciente de las limitaciones que tanto su propia organización como la bioinformática en general tienen y cuya solución se debe acometer para ser realmente útiles y operativas a escala global.

GEO BON

La segunda organización que está intentando operar a escala global es el denominado Grupo de Observación de la Tierra de Observación de la Biodiversidad (GEO BON por las siglas en inglés de Group on Earth Observations on Biodiversity Observation). Es el instrumento de GEO cuya misión es más en concreto la observación de la biodiversidad; se examina también el proceso llevado a cabo por el proyecto (¿es un simple proyecto?) EU BON (¿GEO BON para Europa?). Es cuestionable hasta qué punto difieren ambos, puesto que EU BON parece haber adoptado una dinámica propia mientras que GEO BON no parece ser otra cosa que la aplicación de la informática del GEO a la biodiversidad.

² La Global Biodiversity Informatics Conference (GBIC) es una reunión de todos los investigadores principales, de los políticos y representantes de las principales instituciones de investigación y de las iniciativas en el ámbito de la biodiversidad y de la informática. Su finalidad es desarrollar una visión a largo plazo sobre cómo pueden movilizarse los datos a distintos niveles para hacerlos interoperables y procesables en los ordenadores con distintas finalidades, incluidas: la investigación, las políticas de apoyo, la formación y el hecho de que puedan reutilizarlos los distintos grupos de interés. Se puede consultar más información en <http://www.gbif.org/resource/80771>. (Consultado por última vez el 27 de mayo de 2017).

Aplicación a una escala inferior a la global: dos modelos continentales

Sin embargo, los esfuerzos que parecen estar siendo más operativos suelen darse a una escala inferior a la global, bien nacional, bien supranacional, en especial en países que abarcan un ámbito geográfico continental o casi continental.

En concreto, existe un ejemplo de e-infraestructura de biodiversidad que, constituida en origen por seis Estados miembros de la UE está llamada, en función de el éxito o fracaso que acarree en los próximos años, a constituirse en la infraestructura por excelencia no solo de datos de biodiversidad, sino de análisis (*analytics*), incluso predictivo, de la interrelación de los datos abióticos (químicos, físicos, económicos, etc.) con los de la biodiversidad para crear productos. Estos serán típicamente programas de *software* o de *middleware*, basados en las redes europeas de supercomputación como soporte *hardware*, es decir, aplicaciones y, especialmente, *virtual research environments* (VRE) que permitan modelizar y predecir el estado de la biodiversidad, sus amenazas y posibles soluciones. Se trata de la e-Science and Technology European Research Infrastructure Consortium for Biodiversity and Ecosystem Research —conocida como LifeWatch ERIC—, consorcio constituido el 9 de mayo de 2017 una vez regulada su constitución por decisión de la Comisión Europea y publicada en el Diario Oficial de la Unión Europea el 22 de marzo de 2017.

El modelo LifeWatch se estudia por contraposición, precisamente, de los dos modelos de ciberinfraestructuras que, creadas ambas en Estados Unidos, le han precedido en algunos años en cuanto a sus posibles funciones: por un lado, la Data Observation Network for Earth (Data ONE) y, por otro, aunque añadiendo la función de repositorio de datos mediante una red de sensores, la *National Ecological Observatory Network* (NEON).

Ambos modelos difieren un poco en el diseño como infraestructura debido a que el primero, en el cual se acaba de decidir el traspaso de su entramado institucional y operativo a la Universidad de California dentro de unos años, es un proyecto financiado por la National Science Foundation, mientras que el segundo es un proyecto que deriva de las potestades de gasto del propio gobierno federal y, más en concreto, de la propia Oficina del Presidente en tiempos de la Administración Obama; a pesar de las dudas que surgieron en relación con la Administración Trump, todavía no se han producido cambios.

Aunque la presente tesis no se adentra en ello, sí se describen muy brevemente los otros modelos continentales actualmente existentes, a los que se hará una breve referencia: el South African Biodiversity Institute (SANBI); el Centro de Referencia en Informação Ambiental (CRIA), de Brasil; The Atlas of Living Australia (ALA) o las e-infraestructuras puestas en marcha por la Academia China de Ciencias (CAS).

Coordinación para la realización de análisis globales en la escala espacial y temporal

Hecha la reflexión sobre estos instrumentos, surgió otra cuestión en el seno de la investigación para esta tesis. ¿Son dichos instrumentos redundantes, es decir, utilizan los mismos datos y producen las mismas aplicaciones o VRE (al fin y al cabo, todos se centran en la conservación de la biodiversidad) o bien complementarios (no solo pueden coexistir, sino que es recomendable porque no abordan la biodiversidad de la misma manera)? De hecho, ¿pueden considerarse interdependientes?

Esta última cuestión ha intentado ser abordada precisamente mediante dos proyectos de la Comisión Europea en los que la Universidad de Alcalá ha tenido un papel predominante.

CREaTIVE-B

Uno de estos proyectos es el proyecto del VII Programa Marco de Investigación de la Comisión Europea denominado «Coordination of Research e-Infrastructures Activities Toward an International Virtual Environment for Biodiversity» (CREaTIVE-B).

Este programa analiza en detalle su significado en un caso de estudio, en el que se incluyen organizaciones no solo europeas y, por lo tanto, no es un análisis estrictamente del impacto en la Unión Europea. En efecto, reúne *inputs* de distintos modelos de infraestructuras continentales (LifeWatch, Data ONE, ALA, SANBI, CRIA, CAS). También recurre a infraestructuras de ámbito más global (GBIF). Este proyecto tuvo como objetivo la reutilización de los datos. Identifica el tipo de barreras que obstaculizan no solo el acceso, sino también la reutilización de datos entre las distintas infraestructuras. El uso de un caso de estudio permite analizar en detalle el tipo de barreras: técnicas, semánticas y jurídicas, y las maneras de superarlas. Todo ello quedó plasmado en una Hoja de Ruta que se analiza detalladamente (Alonso García *et al.*, 2014).

GLOBIS-B Terminado este proyecto, y tomando como base sus conclusiones, la Comisión Europea, como parte del Programa Marco de Investigación de Horizonte 2020 (el programa que sustituyó a la serie previa de programas I a VII), puso en marcha el «GLOBal Infrastructures for Supporting Biodiversity research» (GLOBIS-B). Este proyecto, más allá de analizar los pasos que deberían adoptar las infraestructuras para progresar hacia la interoperabilidad global, intenta analizar, desde el punto de vista ya totalmente práctico, unos de los indicadores más importantes que se manejan a nivel científico por GEO BON, el CDB, IPBES y la denominada BIP (Biodiversity Indicators Partnership): las denominadas variables

esenciales de biodiversidad (EBV, por las siglas en inglés de *essential biodiversity variables*).

En este proyecto participan los mismos que fueron protagonistas del proyecto CReATIVE-B y algunas otras infraestructuras, en concreto, las que se listan en la figura que está a continuación.

Acronym	Organisations	Geographic Scope	Website
<i>Project Partners</i>			
UvA	University of Amsterdam (Institute for Biodiversity and Ecosystem Dynamics)	Netherlands	http://ibed.uva.nl/
CU	Cardif University (School of Computer Science and Informatics)	UK	http://www.cs.cf.ac.uk/
GNUBILA	gnúbila France	France	https://gnubila.fr/
CNR	Consiglio Nazionale delle Ricerche (institute of Biomembranes and Bioenergetics)	Italy	http://www.cnr.it/sitocnr/home.html
FI-UAH	Universidad de Alcalá (Instituto Benjamin Franklin)	Spain	http://www.institutofranklin.net
MLU	Martin-Luther-Universität Halle-Wittenberg (German Centre Integrative Biodiversity Research i-Div)	Germany	http://www.idiv-biodiversity.de/idiv/research/geo-bon/
<i>Supporting research infrastructures</i>			
Atlas	Atlas of Living Australia	Australia	http://www.ala.org.au/
BC-CAS	Biodiversity Committee of the Chinese Academy of Sciences	China	http://kepingma.com/index.html
CRIA	Brazilian Reference Center on Environmental Information	Brazil	http://cria.org.br/
DataONE	Data Observation Network Observation	USA	http://dataone.org/
ELIXIR	European Infrastructure for biological information	Europe	http://elixir-europe.org/
GBIF	Global Biodiversity Information Facility	Global	http://gbif.org/
GEOBON	Group on Earth Observations Biodiversity Observation Network	Global	http://geobon.org
GBoWS	Germaplasm Bank of Wild species at Kuming Institute of Botary	China	http://english.kib.cas.cn/
LifeWatch	European Infrastructure for Biodiversity and Ecosystems Research	Europe	http://lifewatch.eu/
NEON	National ecological Observatory Network	Global	http://www.neonine.org/
SANBI	South African National Biodiversity Institute	South Africa	http://sanbi.org/
WCCM	World Data Center of Microorganisms and WFCC-MIRCEN	Global	http://www.wdcm.org/

Figura 3. Infraestructuras de CReATIVE-B. Fuente: Alonso García, 2014.

Hasta la fecha se han ido articulando los mecanismos necesarios para poder calcular dos de las EBV aparentemente más simples y sobre las cuales existen múltiples datos: la de poblaciones totales de especies (*species populations*) y la de rasgos de las especies (*species traits*); queda pendiente todavía la averiguación de la

interoperabilidad necesaria para poder calcular, con las bases de datos actualmente existentes, las interacciones entre especies (*species interactions*).

Para ello, naturalmente, será necesario acometer preguntas tales como: ¿qué son las variables esenciales de biodiversidad? ¿Cómo se definen y de qué manera se representan? ¿Qué es lo que persiguen medir?

Conclusiones

Para terminar, en el cuarto y último apartado, se recoge una serie de conclusiones finales extraídas de cada una de las partes recién expuestas.

BLOQUE I. MARCOS PREVIOS DE CONTEXTUALIZACIÓN

Capítulo 1. Primer marco de contextualización: la revolución digital

1. LA ECONOMÍA DEL CONOCIMIENTO: LA REVOLUCIÓN DIGITAL (DESDE MEDIADOS DEL SIGLO XX HASTA LA ACTUALIDAD)

El punto de partida de este análisis no puede ser otro que la contextualización de la nueva economía del conocimiento, de la revolución digital y los cambios que esto ha supuesto en todos los ámbitos.

1.1. Los orígenes de Internet

En los años sesenta, en plena Guerra Fría, el Departamento de Defensa americano (DoD) llegó a la conclusión de que su sistema de comunicaciones era muy vulnerable y, por lo tanto, decidió desarrollar un proyecto de investigación en redes de conmutación de paquetes.

Hasta entonces, las comunicaciones se habían desarrollado usando la red telefónica conmutada (Ayala, s. f), es decir, la tecnología denominada de conmutación de circuitos,³ con enlaces únicos y en número limitado, entre importantes nodos o centrales. El riesgo fundamental que llevaba aparejada la articulación de las comunicaciones de esta manera era que, en caso de un ataque militar (ruso) sobre esas arterias de comunicación, pudiera quedar aislada una parte del país.

Con la finalidad de mitigar esos riesgos, el Departamento de Defensa, a través de su Agencia de Proyectos de Investigación Avanzados (ARPA) decidió estimular la creación de redes de ordenadores mediante becas y ayudas a departamentos de informática de una multitud de universidades. Esta investigación llevó a la creación de una red experimental de cuatro nodos. Esta red se puso en funcionamiento a finales de 1969 y recibió el nombre de ARPANET.

El objetivo fundamental de esta red era conseguir que la información llegase a su destino, incluso en el supuesto de que parte de la red estuviera destruida, y permitir,

³ Un circuito es una conexión entre llamante y llamado.

con ello, preservar la accesibilidad a la información militar desde cualquier punto del país.

Evolución de la red telefónica

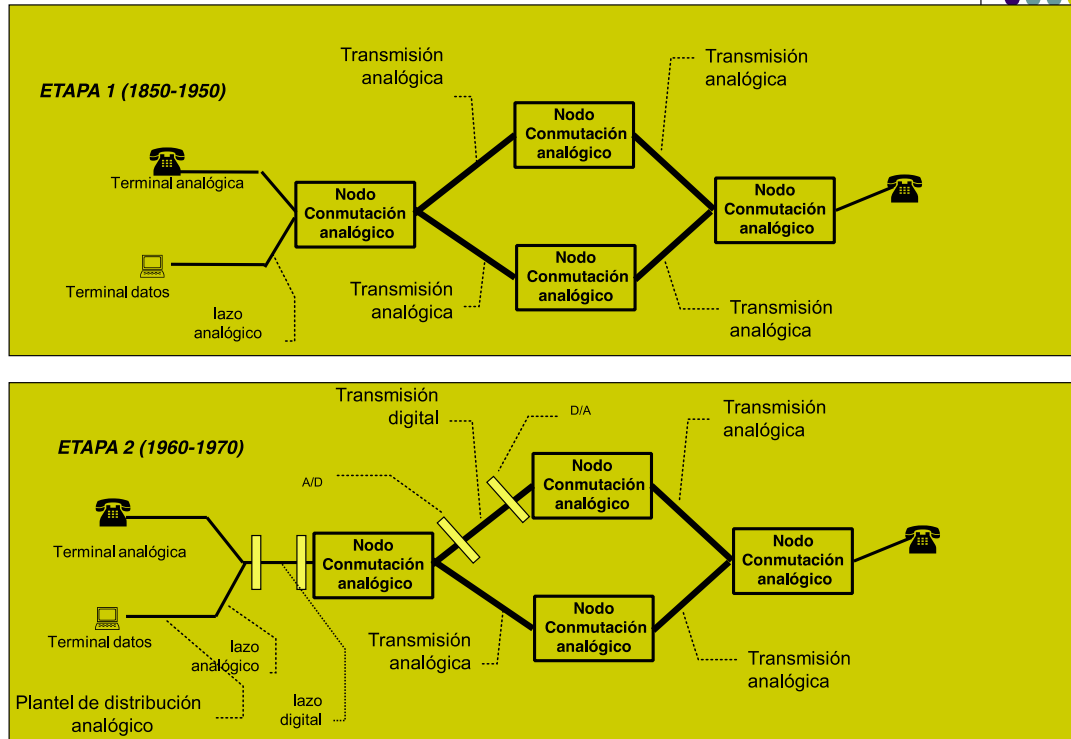


Figura 4. Evolución de la red telefónica. Fuente: Ayala Mojica, s.f.

ARPA desarrolló una nueva tecnología denominada «conmutación de paquetes», cuya principal característica residía en fragmentar la información, dividirla en porciones de una determinada longitud conocidas con el nombre de paquetes.

Cada uno de estos paquetes constaba de una cabecera con la información suficiente para que, cada vez que llegase a un nodo, el paquete fuera encaminado a su destino. Sin embargo, el camino, es decir, la vía para llegar a su destino no estaba prefijado. Cada nodo podía leerla al pasar el paquete por él. De manera que, si una parte de la red cayera o fuese destruida, el flujo de paquetes sería automáticamente redirigido a través de otros nodos alternativos. Los códigos de comprobación permitían conocer la pérdida o corrupción de paquetes, estableciéndose un mecanismo para la recuperación de los mismos.

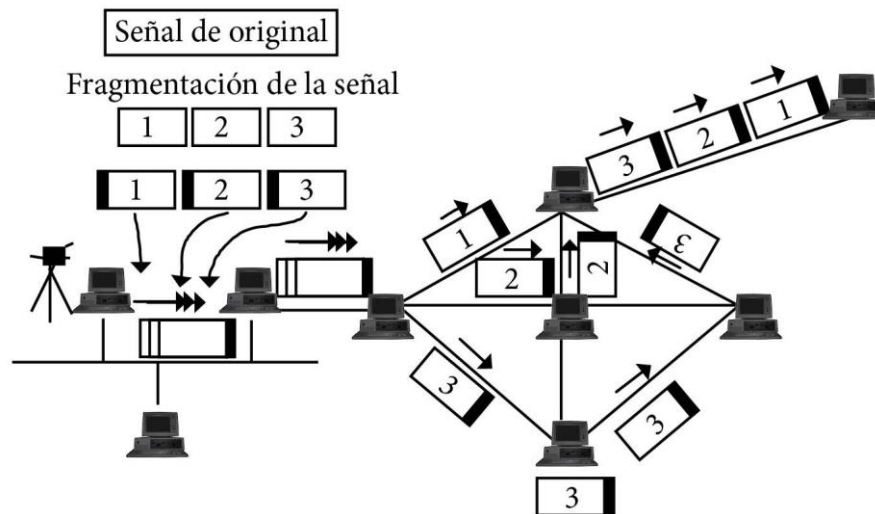


Figura 5. Principio de computación de paquetes. Fuente: Cardozo, s. f.

Este sistema de transmisión reunía múltiples ventajas:

- *La fiabilidad*, con independencia de la calidad de líneas utilizadas y de las caídas de la red.
- *La simplificación de la distribución* de los datos, dado que al contener cada paquete la información necesaria para llegar a su destino, paquetes con distinto objetivo podían compartir un mismo canal o camino de comunicaciones.
- La posibilidad de usar *técnicas de compresión*, que aumentasen la capacidad de transmisión, y de *encriptado*, que permitiesen una codificación. Esto aseguraba la confidencialidad de los datos.

ARPANET continuó creciendo y abriéndose siempre con fines académicos o de investigación. De esta forma, las funciones militares quedaron desligadas de ARPANET y dieron lugar a la creación de una red propia: MILNET.

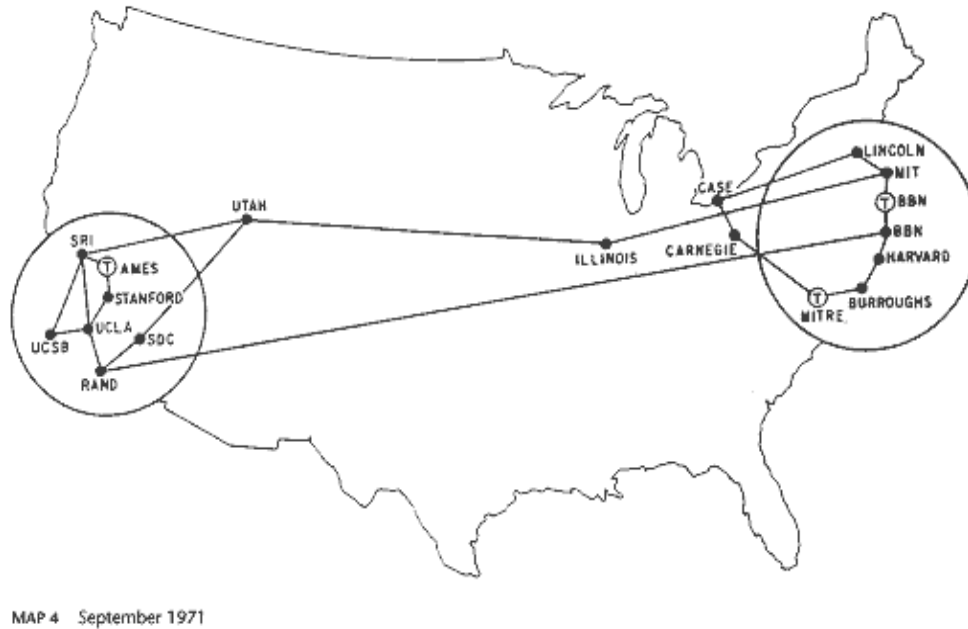


Figura 6. Mapa geográfico de ARPANET de septiembre de 1971. Fuente: Fernández Muñoz, s. f.

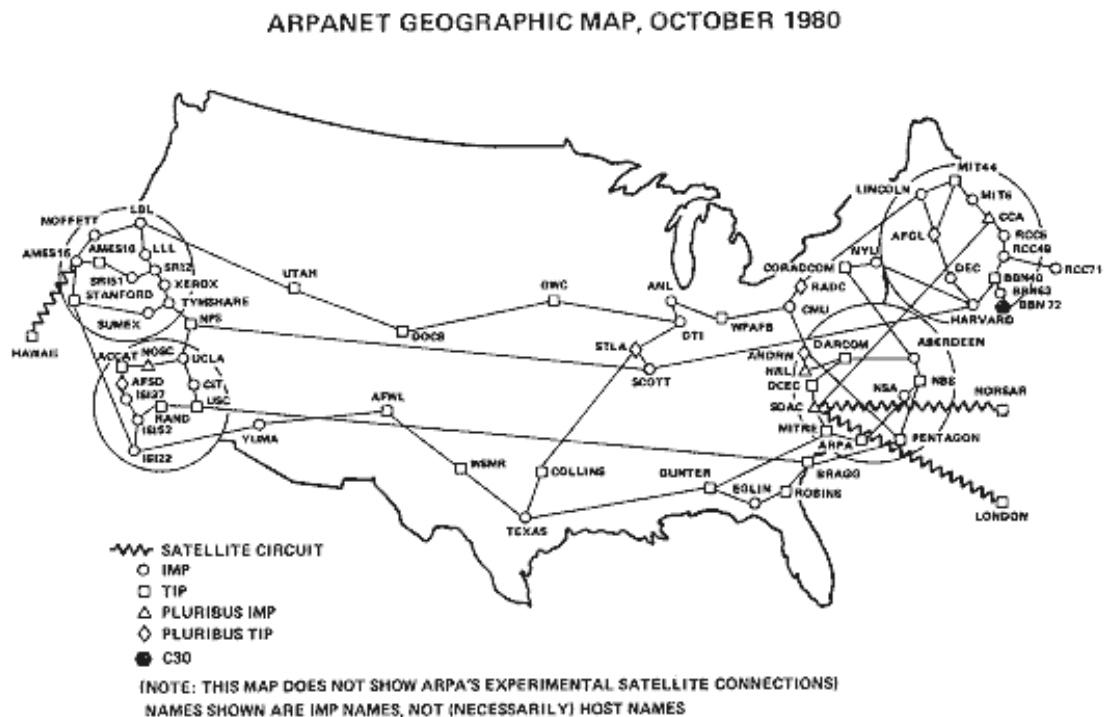


Figura 7. Mapa geográfico de ARPANET de octubre 1980. Fuente: Fernández Muñoz, s. f.

A mediados de los años ochenta, la National Science Foundation estadounidense creó su propia red informática, absorbiendo ARPANET y dando lugar a la creación de una gran red con fines científicos y académicos.

Se inicia, con ello, un gran desarrollo de las redes. Como consecuencia del vertiginoso crecimiento, el protocolo de conexión usado al inicio quedó rápidamente obsoleto y se hizo necesario encontrar un nuevo protocolo de comunicación, como se verá después.

Nació, entonces, el protocolo TCP/IP, un sistema independiente de intercambio de datos entre ordenadores y redes locales de distinto origen, que conservaba las ventajas de la técnica de conmutación de paquetes, estándar y utilizado todavía hoy.

Esto da lugar a lo que hoy se conoce como Internet. En los años ochenta, aunque Internet era una tecnología establecida, solo unos pocos la conocían. Se trataba de una red básicamente de texto.

En los años noventa, en el Centro Europeo de Investigaciones Nucleares (CERN), Tim Berners-Lee retomó una idea de Ted Nelson de usar hipervínculos y desarrolló un proyecto de almacenamiento y recuperación de datos. Se empieza a desarrollar lo que se denomina la World Wide Web o telaraña mundial, es decir, la red de redes.

El científico tenía en la cabeza generar un único espacio, accesible desde cualquier lugar del mundo a través de Internet, que permitiera almacenar, consultar, actualizar y compartir información en forma masiva. Este nuevo sistema permitía vincular información, de forma lógica, a través de las redes. En su origen, la red e Internet no eran términos sinónimos. La red era un paso más, la conectividad de lo que existía en Internet mediante hipervínculos.

En el libro que Berners-Lee publicó sobre el tema, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide* (2000), ya se habla de la forma en la que concibió la red, de su visión, y dice:

La visión que tenía de la red era la de que casi cualquier cosa pudiera estar conectada con otra. Es una visión que nos aporta una nueva libertad y nos permite crecer mucho más rápidamente de lo que nunca hubiéramos imaginado estando encadenados como estábamos con los sistemas jerarquizados de clasificación a los que estábamos ligados.

Esta pequeña contextualización permite extraer algunos de los principios que se encuentran en el origen de la revolución digital. En primer lugar, el hecho de que la red de redes la impulse la National Science Foundation no es ni casualidad ni irrelevante. En efecto, la National Science Foundation es una agencia federal independiente que el Congreso creó en 1950 con la finalidad de «promover el progreso de la ciencia, de fomentar la salud nacional, la prosperidad y el bienestar y garantizar la seguridad nacional» (National Science Foundation, 2016). Por lo tanto, su misión fundamental es apoyar la investigación en todos los ámbitos (ingeniería y

ciencias), por lo que reconoce los importantes beneficios que esta supone para la humanidad.

En segundo lugar, el que la red de redes nazca en Estados Unidos es fundamental, en la medida en que condiciona los principios que se le aplican. En efecto, la red nace como un espacio en el que debe regir la libertad (Akbarzadeh, 2014; Knappenberger, 2014), así que no está sujeto a las normas territoriales del Estado, sino que más bien tiene vocación de quedar sujeto a la autorregulación, como se desprende también de la obra del creador (Berners-Lee, 2000).

Aunque con el transcurso del tiempo se va matizando, sobre todo en lo que respecta a la regulación o sujeción a derecho de lo que exista en la red, sigue latente en el proceso. De alguna forma, condiciona o contribuye a la interpretación de las normas aplicables.

Internet nació con una clara vocación de compartir información y conocimiento, y facilitar, de esta manera, el desarrollo de proyectos y la investigación. Con la aparición de los hipervínculos, la capacidad de ordenar y de estructurar la información, el conocimiento se multiplica exponencialmente hasta el punto de que hoy no es suficiente con «estar» en la red, sino que hay que conseguir, además, que la información que existe sea localizable, esté ordenada, etc.

Esa es la razón por la que surgen muchas de las cuestiones que se van a abordar en este trabajo, ejemplo de ello podría ser la economía colaborativa frente a la economía de «propiedad». Este cambio conceptual supone un cambio mucho más profundo de lo que pueda parecer y está generando muchas tensiones a distintos niveles.

1.2. El concepto de Internet

Aunque responder a la pregunta de qué es Internet pueda parecer obvio, no lo es tanto. El término «Internet» tiene toda una serie de acepciones que García Mexía (2009) sistematiza en dos grandes perspectivas: una circunstancial y otra antropológica.

En lo que respecta a la perspectiva circunstancial, Internet ha venido a desafiar los términos «tiempo» y «espacio» (García Mexía, 2009). En efecto, Internet ha modificado la percepción del tiempo dando lugar al nacimiento del término «tiempo real» para poner de manifiesto la inmediatez de los actos, por ejemplo: el envío de un correo electrónico.

En lo que respecta al espacio, se ha creado un espacio virtual conocido con el nombre de «ciberespacio»: espacio indefinido no asociado a ningún territorio o superpuesto a todos los territorios, según se mire. En cualquier caso, resulta difícil

de ligar con el concepto «territorial». De hecho, el espacio virtual, decía el creador de la World Wide Web (Berners-Lee, 2000), tenía una vocación de espacio global, no sujeto a ninguna jurisdicción.

La perspectiva antropológica queda dividida en cinco dimensiones: la política, la social, la económica, la cultural y la tecnológica (García Mexía, 2009).

La primera dimensión, la *política*, afecta al gobierno de la red, sobre todo habida cuenta del despliegue mundial y la crecientemente masiva accesibilidad. No puede negarse que la revolución tecnológica que se está produciendo está teniendo impacto en la forma en la que se ejercer el poder. Además, la aparición de Internet está transformando las tradicionales relaciones entre Estado y sociedad, haciendo de esta una sociedad de la información y del conocimiento (COM, 2007).

Más allá de eso, se ha acuñado ya el término de «democracia digital» —conocida en inglés como *e-democracy*— que es el uso de las TIC (informática, Internet, computación móvil y telecomunicaciones) para crear espacios de diálogo y reflexión social, acceso a información de sujetos políticos (planes de trabajo, ideología), ejercicio de los derechos de participación o políticos y la mejora de los procesos electorales en la relación entre ciudadanos, sujetos políticos e instituciones electorales.

La democracia usa las nuevas tecnologías de la información y los medios alternativos de comunicación para satisfacer sus necesidades en beneficio de todos y para mejorar procesos dentro de una república democrática o democracia representativa. A pesar de ello, es un desarrollo político aún en su infancia, y objeto de muchos debates y actividad dentro de gobiernos, grupos cívicos y sociedades en todo el mundo.

Es un concepto todavía poco elaborado, que, sin lugar a dudas, está teniendo un impacto incuestionable, como demuestran ejemplos muy actuales. Uno de ellos podría ser el reciente intento de golpe de Estado ocurrido en Turquía, en el que fueron clave los medios digitales. Más allá de eso, el presidente del país usó dichos canales digitales como una manera de parar el golpe de Estado. De igual manera, pueden citarse los ejemplos de uso de las redes sociales por parte del grupo Anonymous, para manifestarse de forma masiva, o el ejemplo de Aaron Swartz (Knappenberger, 2014). Es un proceso que acaba de empezar.

Una segunda dimensión sería la *social*, en la medida en que en la red nacen y se desarrollan relaciones sociales de toda índole. Se hacen todo tipo de intercambios de información. Constituye un sistema, habida cuenta de que sus terminales están interconectadas y pueden llegar a ser interdependientes. Muestra de esto es el crecimiento de las redes sociales y la multiplicación de intercambio de información a través de ellas. De hecho, está cambiando la forma en la que las personas se

relacionan. Sobre las cuestiones que plantea la dimensión social, se volverá más adelante por las implicaciones que a los efectos de este trabajo tiene trabajar sobre las necesarias interdependencias.

La tercera dimensión sería la *económica*, que, de hecho, ha llevado a hablar ya de la economía digital. Como ya puso de manifiesto la OCDE (OECD, 2015), las tecnologías de la información y del conocimiento (TIC) desempeñan una función cada vez más trascendente en la economía pues son objeto de una gran parte de la inversión. Contribuyen, además, notablemente al incremento de la producción y de la productividad. De hecho, en la actualidad todas las tendencias en los mercados están llevando a las empresas a la transformación digital. La duda es cuáles van a ser capaces de integrar las nuevas formas de trabajo: automatización, transformación digital, etc.

La cuarta dimensión, la *cultural*, está resultando, sin lugar a dudas, revolucionaria. Desde este punto de vista, Internet está siendo el principal motor de la sociedad de la información y del conocimiento. Las bibliotecas están creando repositorios digitales enormes y el procesamiento del conocimiento se está multiplicando exponencialmente. La clave, ahora, es ser capaces de poner en relación todo el conocimiento, trabajar sobre la forma de gestionarlo, etc. Esta perspectiva cultural es de las más relevantes a los efectos de este trabajo, como se irá viendo.

Por último, no puede ignorarse, la *dimensión tecnológica*, pues, a fin de cuentas, Internet se define como una red de ordenadores. No cabe duda de la relevancia que tiene la dimensión tecnológica, que ha llevado a algunos autores a hablar de los tres estratos de Internet (Benkler, 2000; Lessig, 2002).

En primer lugar, el estrato *físico* de Internet está constituido por la propia red, es decir el complejo entramado de conexiones. De este estrato físico formaría también parte el ordenador personal.

En segundo lugar, el segundo estrato, el *contenido*, está integrado por una multiplicidad de fuentes de información y de conocimiento que Internet pone a disposición de los usuarios.

En tercer lugar, el estrato *lógico*, plasmado en la denominada interfaz, más concretamente la conexión física y funcional entre los ordenadores y las redes de ordenadores, posible gracias al *software*, los estándares de comunicación y los protocolos específicamente diseñados.

En este trabajo, se abordarán sobre todo cuestiones relacionadas con el contenido y las fuentes de información, y de modo, menos profundo las cuestiones relacionadas con los estándares, etc.

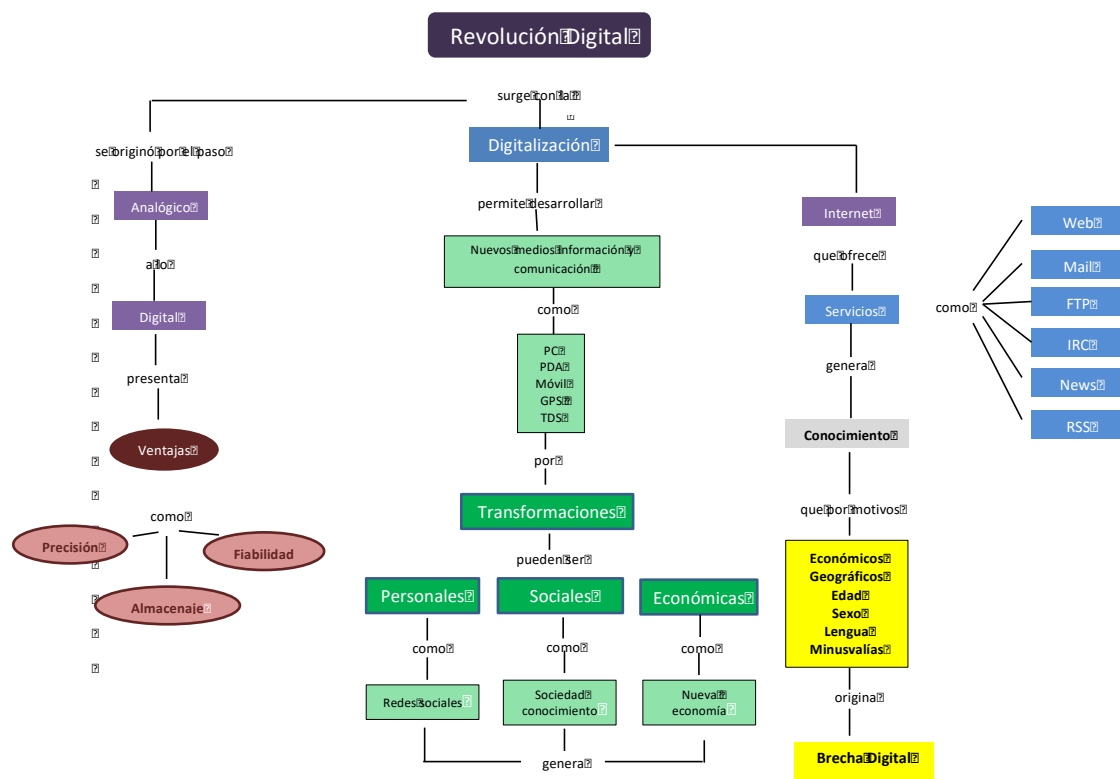
Como se verá a lo largo de este trabajo, esta introducción de las perspectivas, dimensiones y estratos marca la gran revolución que está suponiendo (¿o ya ha supuesto?) Internet.

Casi puede decirse que ya se ha producido en el ámbito físico, aunque queda mucho por trabajar en lo que respecta a la eliminación de barreras y trabas (legales, técnicas, etc.) para que lo que se está produciendo en el estrato físico tenga su reconocimiento en todos los ámbitos (legal, científico, etc.).

1.3. La gestión del conocimiento en la economía del conocimiento

El cambio que Internet ha supuesto es tan grande que se habla de revolución digital e, incluso, de *homo digitalis* (Rodríguez, 2005);⁴ el *homo sapiens* parece estar ya en proceso de desaparición. Se habla ya de la «singularidad», un paso más allá del *homo digitalis*, la convivencia de hombres y robots... Las fechas que se barajan para alcanzar una inteligencia artificial que pueda convivir con el hombre son entre 2029 y 2045 (Cordeiro, 2015).

Una representación conceptual de lo que es la revolución digital (García Martínez y Turégano, 2010) puede servir de punto de partida del análisis:



⁴ Uno de los primeros autores en hablar del *homo digitalis*. Después, varios autores retoman el término.

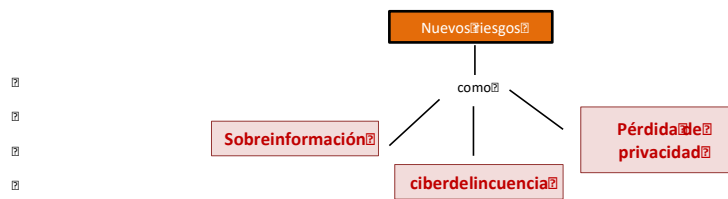


Figura 8. Esquema conceptual: revolución digital. Fuente: Martínez Navaro y Turégano García.

Ya a principios de la década del 2000, el profesor H. Stroz (Mokyr, 2002) publica un interesante estudio sobre esta «última» revolución en la que está ahora mismo inmersa la sociedad.

Observa el impacto que esto ha tenido en el desarrollo de las personas. En los primeros siglos, el conocimiento solo estaba al alcance de unos pocos y, por lo tanto, «podía desaparecer», «perderse». De hecho, esta es una de las razones que pueden explicar «algunos de los pasos atrás» de la historia de la humanidad.

En efecto, el conocimiento que sirvió de base para el crecimiento de algunas culturas, no se conservó y, por ende, no pudo servir de punto de arranque, de referencia, para las siguientes.

Con la revolución del conocimiento, se ha multiplicado el mismo de forma exponencial. El problema hoy no es la cantidad de conocimiento (o tal vez sí, puesto que ya se habla de sobreinformación), sino el acceso, la diseminación del mismo y la estructuración para que el conocimiento existente sirva para seguir construyendo conocimiento nuevo para avanzar.

Algunos autores hablan ya de la economía del conocimiento (David y Foray, 2003). El nacimiento de esta economía tiene origen en una multiplicidad de factores, una de ellas es, sin duda, la aceleración del conocimiento, como ya mencionó con anterioridad Mokyr (2002). En esa misma línea, hoy la innovación es una de las claves a todos los niveles, hasta tal punto que casi puede decirse que de la capacidad de innovación depende la supervivencia.

Una de las claves a la hora de analizar esta idea es la importancia de diferenciar entre información y conocimiento (David y Foray, 2003). El conocimiento podría considerarse una forma de empoderar al que lo tiene con la capacidad de realizar una acción física o intelectual. Por lo tanto, es, prosiguen los autores, una cuestión de capacidad cognitiva. La información toma la forma de datos estructurados y ordenados, pasivos e inertes hasta el momento en que la usan aquellos que tienen el conocimiento necesario para interpretarla y procesarla.

La diferencia es clara a la hora de hacer una reproducción. Mientras la reproducción de la información solo cuesta lo que cuesta una copia (hoy en día, una cantidad que

tiende a cero gracias a las nuevas tecnologías), reproducir el conocimiento es un proceso mucho más costoso habida cuenta de que algunas capacidades cognitivas no son tan sencillas de replicar expresamente, ni de transferir a otros.

Ahondando en las cuestiones que surgen en relación con el conocimiento, se plantean, de inicio, dos cuestiones fundamentales que procede elaborar. La primera es la capacidad de «guardar», de almacenar el conocimiento de una forma accesible y, la segunda, posiblemente consecuencia de la anterior, es que el conocimiento no quede enterrado por el aluvión de datos que existen al alcance de todos. Se hace necesario filtrar toda esa información e identificar la relevante.

Es clave a los efectos de esta revolución en que la sociedad está inmersa ver de qué manera las nuevas tecnologías van a permitir mejorar la capacidad de resolver cuestiones complejas (Nielsen, 2012). Este autor pone ejemplos diversos que muestran de qué manera, hoy, la utilización de la tecnología permite «buscar» y «localizar» microexpertos de distintos temas. En efecto, la red permite hoy acceder con facilidad al conocimiento directamente o a la persona que tiene dicho conocimiento.

Lo importante o relevante en la sociedad del conocimiento no es ya el conocimiento individual, sino el colectivo (Nielsen, 2012). El paso del conocimiento individual al conocimiento colectivo plantea toda una serie de cuestiones y de cambios. El conocimiento es un bien público en el sentido de que el consumo de uno no reduce el de los otros (Nielsen, 2012). Si bien los costes privados de su adquisición son importantes en términos de esfuerzo y, a menudo, de recursos reales. En efecto, el conocimiento colectivo no solo no excluye el conocimiento individual, sino que contribuye a alimentarlo. De alguna manera, el «incremento de conocimiento» colectivo contribuye a un desarrollo del conocimiento individual —por supuesto, siempre que ese conocimiento colectivo, sea accesible—.

Es evidente que las herramientas para gestionar el conocimiento no son las mismas en esta sociedad del conocimiento. Además, los problemas que plantea el acceso a esta información son diferentes.

La realización de este análisis arranca de la sociedad del conocimiento y plantea una serie de cuestiones que arrancan de tres fundamentales sobre las que se replantea el cómo «estructurar» el conocimiento hoy, en un entorno tan tecnológico y que se desarrollan a continuación (Edwards *et al.*, 2012).

La primera cuestión es de qué manera están cambiando las infraestructuras de conocimiento. La segunda, de qué manera los cambios en las infraestructuras de conocimiento refuerzan o redistribuyen la autoridad, la influencia y el poder y, la tercera, cuál es la mejor forma para estudiar, saber e imaginar de qué manera las infraestructuras del conocimiento se van a construir de ahora en adelante.

Ligadas a estas cuestiones fundamentales, se plantean otras tantas, por ejemplo: ¿Cómo puede asegurarse la pervivencia de los datos a largo plazo? ¿Cómo asegurar la calidad? ¿Cómo gestionar esta cantidad de datos? Son cuestiones tanto de carácter físico —ligadas a la potencia de los ordenadores y capacidad de gestión— como de carácter filosófico, como se irá analizando a lo largo del presente trabajo de investigación.

Se requiere un nuevo enfoque para comprender las transformaciones que se están produciendo en la forma en la que las personas crean, comparten y cuestionan el conocimiento. En efecto, como ya se ha puesto de manifiesto, la red nace como una red de redes cuya finalidad, ya desde sus inicios, es poner en común, sobre todo desde el punto de vista del contenido de carácter científico.

A través de las nuevas tecnologías, en este momento se están replicando en la red las infraestructuras que existen fuera de ella: bibliotecas virtuales, acceso a los datos de los museos, enciclopedias en línea, etc.

Aunque este hecho es válido como punto de arranque, probablemente será necesario hacer cambios más profundos. Igual que ha ocurrido siempre, los usuarios de conocimiento pasan de ser «consumidores pasivos» de contenidos a participar activamente en su creación, desarrollo y difusión en un entorno colaborativo. Lo que sí ha supuesto un cambio es la intensidad y la capacidad de desarrollar contenidos, que se han incrementado. Es curioso de qué manera, hoy, se han difuminado las líneas de separación entre el mundo de los «productores de conocimiento» y el de «los consumidores de conocimiento». Unos y otros se retroalimentan y las vinculaciones entre los distintos temas van proliferando a una velocidad vertiginosa. Hoy, el conocimiento es «colaborativo». Es decir, que sobre cada cosa nueva que se descubre y se conoce se va construyendo, a su vez, más conocimiento. Tiene un efecto multiplicador.

Esto genera la aparición de nuevas herramientas en la red, donde se crean espacios puestos a disposición de los autores en los que compartir conocimiento, artículos y demás informaciones sin coste para los consumidores de conocimiento.

Todos estos importantes cambios en las instituciones del conocimiento tradicionales —bibliotecas, universidades, etc.— están en proceso de reajuste a la nueva realidad de compartir y han decidido o están decidiendo abrir virtualmente «sus puertas», dando acceso al conocimiento que albergaban. La forma de abordar hoy el conocimiento está cambiando, distingue dos modos de acercarse al conocimiento: el tradicional, Modo 1, generado por contexto, disciplina, etc. y el actual, Modo 2, mucho más multidisciplinar, amplio (Gibbons *et al.*, 1994).

Así, la nueva forma de conocimiento es multidisciplinar y mucho más heterogénea en la forma en la que se involucran las personas en producir dicho conocimiento.

Por ello, se han generado nuevas estructuras para abordarlo. Además, una de las fuentes de estímulo de la producción de conocimiento ha sido, como ponen de manifiesto estos autores, hacer frente a los nuevos problemas globales. Por lo tanto, desde el inicio existe una sensibilidad hacia el impacto de la investigación.

Los criterios de control de calidad también están cambiando: no solo se tienen en cuenta criterios científicos, sino también otros de índole cultural, social, etc. Todo se encuentra en un proceso de cambio, de adaptación a esta nueva realidad: la tecnología y las nuevas formas de acceder al conocimiento. El impulso para compartir los datos viene de distintas fuentes.

En algunos casos, lo han impulsado los responsables políticos, conscientes de la necesidad (en Estados Unidos, la National Science Foundation, por ejemplo); en otros, lo están impulsando los «productores de conocimiento» y, también, en otros casos, los investigadores y «propietarios» del conocimiento se están concienciando de la necesidad de compartir conocimiento.

Como se verá más adelante, el acceso al conocimiento es clave y se están produciendo movimientos de eliminación de barreras: políticas de *Open access*, *Free access*, etc.

El informe *Sustainable Economics for a Digital Planet: ensuring long-term access to digital information* (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010) pone de manifiesto lo siguiente:

SUSTAINABLE ECONOMICS FOR A DIGITAL PLANET

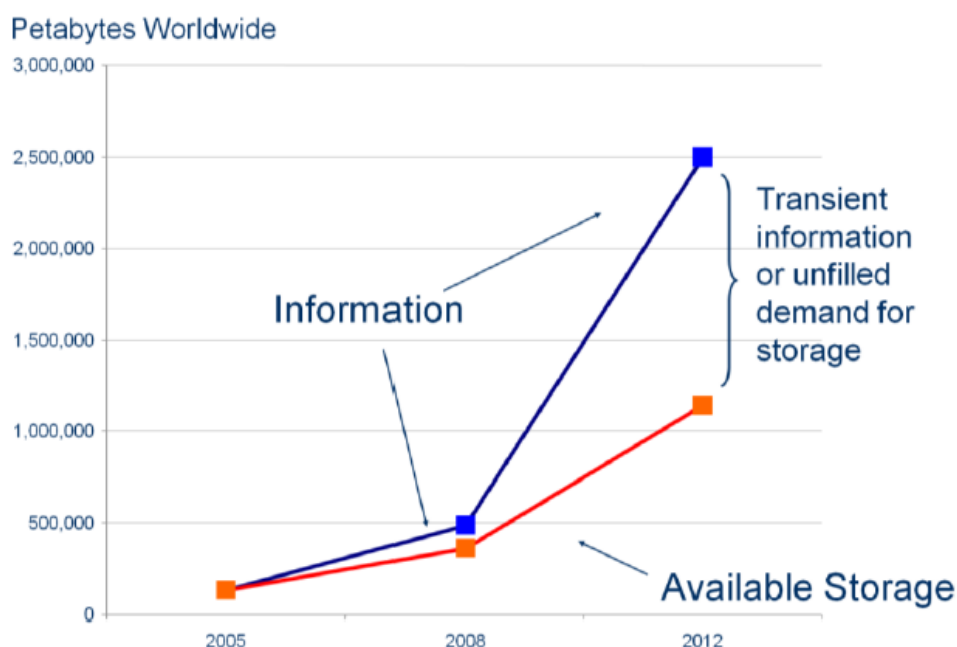


Figura 9. Crecimiento de la información y tendencias del almacenamiento (proyecciones de crecimiento de la tasa de crecimiento de almacenamiento disponible de creación de información global). Fuente: IDC Digital Universe White Paper. Fuente: Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010.

Por eso, algunas de las cuestiones que se plantean son las siguientes: ¿qué información digital es necesario preservar?, ¿quién tiene que hacerlo?, ¿quién va a hacerse cargo desde el punto de vista económico de esta conservación? Estas preguntas que ya se plantearon en el año 2010, siguen teniendo hoy plena vigencia (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010).

Después de este recorrido por la sociedad del conocimiento, la nueva revolución digital y lo que esta ha significado, el siguiente paso será entrar en el análisis de las nuevas formas de gestión del conocimiento.

1.4. Del *data warehouse* al *big data*

Las «antiguas» formas de compartir, almacenar y difundir conocimiento han quedado obsoletas de golpe. Como primeros pasos de esta revolución aparece, a finales del siglo XX, una nueva forma de gestionar los datos que se conoce como *data warehouse* (almacén de datos). De alguna forma, aplica la tecnología a lo que existe y multiplica la potencia de gestión de datos.

Tomando la definición de Bill Inmon, el *data warehouse* se define en términos de las características del repositorio de datos:

- Orientado a *temas*: Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- *Variable* en el tiempo: Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.
- *No volátil*: La información no se modifica ni se elimina, sino que, una vez almacenado un dato, este se convierte en información de solo lectura, y se mantiene para futuras consultas.
- *Integrado*: La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.

El autor propugna una metodología descendente (*top-down*) a la hora de diseñar un almacén de datos, ya que de esta forma se considerarán mejor todos los datos corporativos. En esta metodología los *data marts* fueron creados después de haber terminado el *data warehouse* completo de la organización.

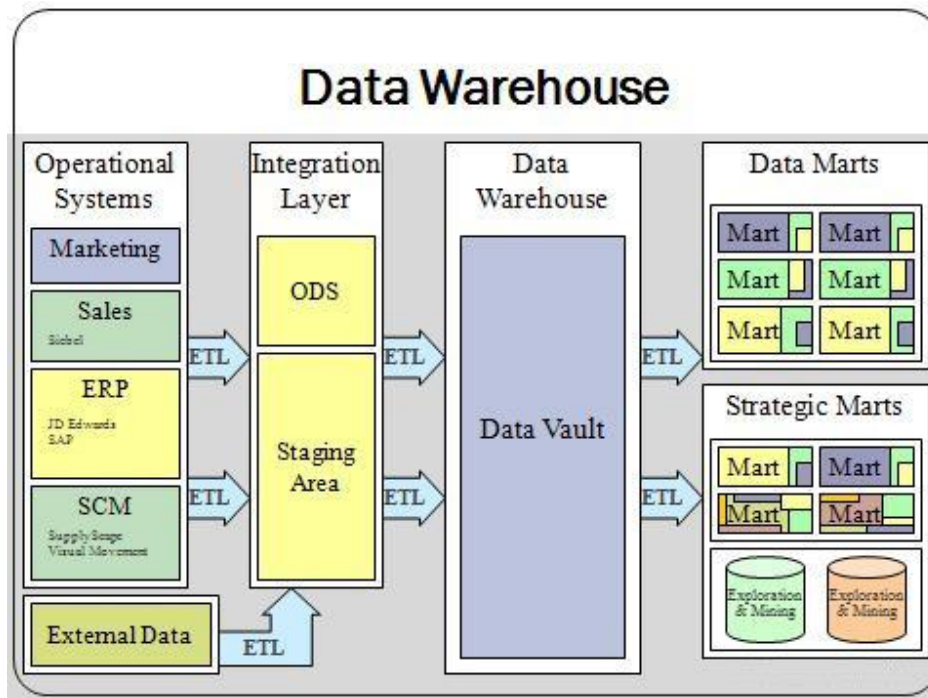


Figura 10. Data Warehouse. Fuente: Wikipedia.

No todos los autores están de acuerdo en esta construcción y, de hecho, Kinball defiende otra construcción del *data warehouse*; sin embargo, queda fuera del ámbito de esta tesis analizar las discusiones con un nivel mayor de profundidad.

Aunque fue la precursora de la inteligencia empresarial y el instrumento para empezar a crear *business analytics* (CRM), hoy ya puede considerarse una tecnología obsoleta. En efecto, la exponencial multiplicación de datos hace que este sistema que trabaja con base en repositorios fijos no sea suficiente.

A resultas del aluvión de datos, se hace patente la necesidad de e-infraestructuras, es decir, infraestructuras tecnológicas que permitan gestionar todos los datos y hacerlos rápidamente accesibles. Toda esta gestión está siendo definida como *big data* (Mayer-Schonberger y Cukier, 2013).

¿Qué es el *big data*? Es un concepto que hace referencia a conjuntos de datos tan grandes que aplicaciones informáticas tradicionales del procesamiento de datos no son suficientes para tratar con ellos y a los procedimientos usados para encontrar patrones repetitivos dentro de esos datos (Big data, s. f.).

Se considera *big data* la gestión y análisis de enormes volúmenes de datos que no pueden ser tratados de manera convencional, ya que superan los límites y capacidades de las herramientas de *software* habitualmente utilizadas para la captura, gestión y procesamiento de datos (López López, 2014).

El objetivo de *big data*, al igual que sucede con los sistemas analíticos convencionales, es convertir el dato en información que facilita la toma de decisiones, incluso en tiempo real. Tal como se dijo al principio, el concepto «tiempo» se ha acelerado.

El dato tiene, hoy, la consideración del nuevo «oro» (McAfee y Brynjolfsson, 2012), es decir, se ha producido una transformación. ¿Qué tiene el *big data* de especial? El movimiento de *big data*, al igual que el movimiento de *analytics* antes, busca extraer inteligencia de los datos y traducirlo en ventajas empresariales.

No obstante, el hecho de asociar el concepto de *big data* a grandes volúmenes de datos no es nuevo. La gran mayoría de las empresas llevan ya mucho tiempo manejando grandes volúmenes de datos y han desarrollado *data warehouses* y potentes herramientas analíticas que les permiten tratar de forma adecuada esos grandes volúmenes. La evolución de la tecnología y los menores costes del almacenamiento han hecho que los volúmenes manejados por estas aplicaciones hayan aumentado de manera muy importante.

Sin embargo, existen tres diferencias clave (McAfee & Brynjolfsson, 2012) con el proceso anterior: la primera diferencia hace referencia al *volumen*. En efecto, desde 2012, se crean cerca de dos millones y medio de exabits cada día y, cada cuarenta meses aproximadamente, este número se multiplica por dos. En la actualidad, hay más datos cruzando Internet cada segundo de los que existían almacenados en toda la red hace veinte años. Esto permite a las compañías trabajar con muchos petabits de datos en un único conjunto de datos, y no solo datos procedentes de Internet. Por ejemplo, se estima que Walmart obtiene más de 2,5 petabits de datos cada hora de las transacciones que realiza con sus clientes. Un petabit equivale a un trillón de bytes. Un exabit es igual a mil veces esta cantidad o a mil millones de gigabits.

La segunda diferencia es la *velocidad*. Para muchas aplicaciones, la velocidad de creación de los datos es incluso más importante que el volumen. La información en tiempo real o casi en tiempo real permite a una compañía ser mucho más ágil que sus competidores. Por ejemplo, un grupo de estudio del MIT Media Lab utilizó la localización de datos a través de los teléfonos móviles para inferir cuántas personas se encontraban en el aparcamiento de Macy's en el *Black Friday*, día que marca el principio de la temporada de compras de Navidad en Estados Unidos. Esto permitió hacer una estimación de las compras de este día incluso antes de que lo hiciera Macy's. Es evidente que tener esta información rápidamente puede representar claras ventajas competitivas para los analistas de Wall Street o para los gestores a pie de calle, por ejemplo.

La tercera diferencia es la *variedad*. El *big data* toma la forma de mensajes, actualizaciones e imágenes que se publican en las redes sociales, que se leen en los

sensores, señales de GPS de los teléfonos móviles y muchas más fuentes de donde pueden obtenerse. Muchas de las fuentes más importantes del *big data* son relativamente nuevas. Las enormes cantidades de información de las redes sociales, por ejemplo, solo tienen la antigüedad que tienen dichas redes sociales —Facebook se lanzó en 2004 y Twitter, en 2006—.

Esto mismo se aplica a los *smartphones* y a otros dispositivos móviles que ahora suministran enormes cantidades de datos relacionados con las personas, actividades y ubicaciones. Teniendo en cuenta que estos dispositivos están en todas partes, es fácil olvidar que el iPhone se presentó en 2004 y el iPad en 2010. Por lo tanto, las bases de datos estructuradas que almacenaban la mayor parte de la información corporativa hasta hace bien poco no están preparadas para almacenar y procesar *big data*. Al mismo tiempo, la disminución paulatina de los costes de todos los elementos computacionales —almacenamiento, memorias, procesadores, ancho de banda y todos los relacionados— hacen que los enfoques intensivos en datos, antes muy caros, están abaratándose por momentos.

En efecto, como señala López López (2014), la información que procesan los *data warehouses* es información estructurada que ha pasado por numerosos filtros de calidad para poder garantizar que la información de salida tiene una precisión y una exactitud determinadas. Sin embargo, al hablarse de *big data*, se hace referencia a información que puede estar semiestructurada o no tener ninguna estructuración.

La gestión de esta información desestructurada precisa de una tecnología diferente y permite tomar decisiones basadas en información que tiene importantes grados de inexactitud. Muchos de estos algoritmos se relacionan con los tratamientos de sistemas avanzados de lógica difusa.

A estas tres cuestiones de base, hay que añadir dos diferencias más: la cuarta, la *veracidad*, que señala López López (2014) hace referencia a la confiabilidad en los datos, la capacidad de extraer datos de calidad eliminando la imprevisibilidad inherente de algunos, como el tiempo, la economía etc., para, de esta forma, llegar a una correcta toma de decisiones

Para terminar, la quinta es el *valor* (López López, 2014). La importancia del dato para el negocio, es decir, que es fundamental saber qué datos son los que se deben analizar. Tanto es así, que ya se empieza a hablar del científico de datos, un profesional con perfil científico, tecnológico y visión de negocio. De hecho, hoy los datos son el foco de atención de la sociedad del conocimiento y de la economía del conocimiento que quiere maximizarlos, rentabilizarlos.

Si las e-infraestructuras ya planteaban muchas cuestiones, la problemática del *big data* las ha acentuado y acelerado.

En los puntos relacionados con la bioinformática y la interoperabilidad, se analizará el impacto del *big data* en lo que respecta a la ciencia y a la biodiversidad en concreto. Ahora se van a analizar con detalle las nuevas formas de gestión del conocimiento del siglo XXI.

2. LA NUEVA FORMA DE GESTIONAR EL CONOCIMIENTO EN EL SIGLO XXI: ¿QUÉ SON LAS E-INFRAESTRUCTURAS?

El término e-infraestructura se acuña para entornos de investigación innovadores y proporciona a los científicos modernos acceso directo a recursos compartidos, distribuidos y heterogéneos (Candela *et al.*, 2009).

Existen muchas definiciones y nombres para referirse a estas infraestructuras (Navarro *et al.*, 2013). Las infraestructuras de innovación se definen como la combinación de elementos que soportan los procesos de innovación y con economías de escala que trascienden las necesidades de las empresas individuales (Moshe y Teubald, 1986).

Ciberinfraestructura es, a fin de cuentas, el conjunto de prácticas organizativas, infraestructuras técnicas y normas sociales que, de forma colectiva, hacen posible el correcto funcionamiento de un trabajo científico a distancia. Los tres elementos son parte del diseño y de la ingeniería, por lo que la ciberinfraestructura será fallida si se ignora cualquiera de los tres (Edwards *et al.*, 2007).

Las ciberinfraestructuras están conformadas por *hardware*, datos y redes, sensores habilitados digitalmente, observatorios e instalaciones experimentales y un conjunto interoperable de *software* y de servicios de *middleware* y herramientas. Es también fundamental realizar inversión en equipos multidisciplinares y profesionales de las ciberinfraestructuras con experiencia en desarrollo de algoritmos, aplicaciones de sistemas y desarrollo de aplicaciones con la finalidad de aprovechar al máximo la potencia de la ciberinfraestructura a efectos de crear, diseminar y preservar los datos científicos, la información y el conocimiento (National Science Foundation, 2007).

Mientras que en Estados Unidos se usa con frecuencia el término «ciberinfraestructura», en Europa es más común el de «e-infraestructura» (Edwards *et al.*, 2009). Como la distinción no es especialmente relevante, a lo largo de este trabajo se usarán ambos términos de forma indistinta.

La infraestructura de conocimiento es la infraestructura física y organizacional que, siendo parte del marco general y produciendo, financiando, coordinando, supervisando y valorando esfuerzos de financiación, sirve de apoyo y es empleada por las empresas para innovar (Doloreux, 2002).

No puede negarse la dificultad que existe a la hora de definir las e-infraestructuras. De hecho, las personas que estudian de qué manera la tecnología afecta a la transformación organizacional reconocen cada vez más su naturaleza dual y paradójica (Star y Ruhleder, 1996). En efecto, esta actúa al mismo tiempo como motor y como barrera, parametrizable y rígido a la vez. Es simultáneamente producto y proceso.

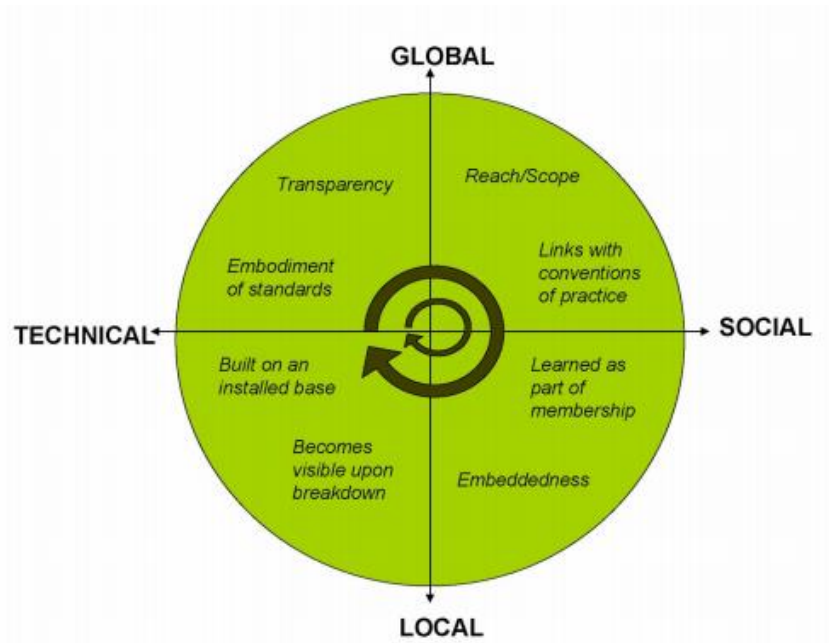


Figura 11. La ciberinfraestructura como una distribución entre los ejes técnico/social y global/local. Fuente: Knobel, 2007.

Esta podría ser una representación de una síntesis de las definiciones (Knobel, 2007).

Durante los últimos años, se ha multiplicado el uso del término «infraestructura». Con carácter general, hace referencia al concepto de gran tamaño de sistemas y servicios que funcionan correctamente a largo plazo (Edwards *et al.*, 2009).

Desde el punto de vista conceptual, las e-infraestructuras pueden representarse en tres capas (Andronico *et al.*, 2011). La primera capa está representada por los instrumentos científicos y los experimentos que proporcionan una ingente cantidad de datos; la segunda, por el diseño de la Red, la red de centros de procesamiento de datos y los *softwares* intermedios como pegamento de los recursos y, la tercera, el nivel más alto, viene representada por los investigadores que realizan sus actividades con independencia de sus localizaciones geográficas, interactúan con los colegas, comparten y acceden a los datos.

Estas capas hacen referencia a lo que se conoce como la «gridificación»⁵ regional y mundial, que, a fin de cuentas, es la multiplicación de los efectos de creación de e-infraestructuras regionales y globales interconectadas (Andronico *et al.*, 2011).

Los datos y los documentos tienen más valor cuando están interconectados que cuando son independientes. El término infraestructuras de la información se usa como un término colectivo que engloba lo técnico, lo social, lo político, las personas, la tecnología, las herramientas y los servicios utilizados para facilitar el uso colaborativo de contenido en el tiempo y en la distancia (Borgman, 2010).

Las ciberinfraestructuras se definen solamente mediante ejemplos, haciendo especial énfasis en su naturaleza integradora, colaborativa y distribuida de las nuevas formas de investigación (Atkins, 2003).

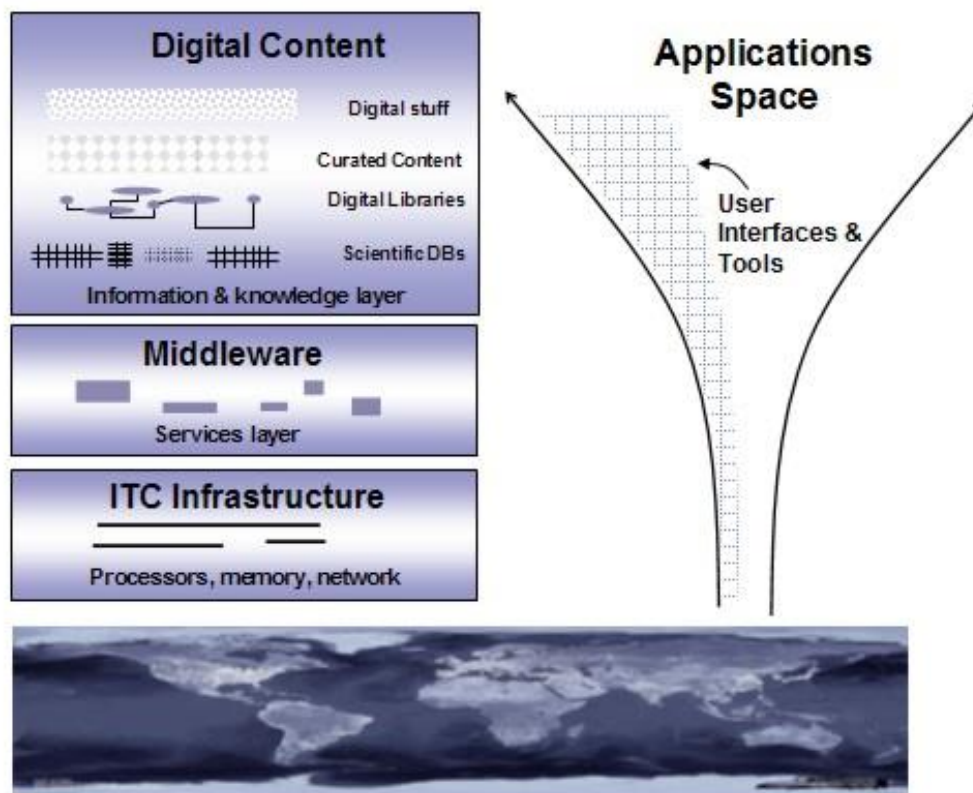


Figura 12. A menudo se ha utilizado un modelo en capas para describir la «infraestructura de la información». En origen, la infraestructura de información (que recibe distintos nombres) incluye el *hardware* informático, las redes y las telecomunicaciones, tarjetas de memoria y el *software* operativo para hacer que estos trabajen conjuntamente. En esta concepción geológica de recursos interconectados de información global y telecomunicaciones, las capas inferiores de los sistemas informáticos y las redes permiten las capas superiores de *middleware*, colecciones de contenido digital y aplicaciones. Esta representación gráfica sugiere que el contenido digital en sí mismo

⁵ Por la referencia a la palabra inglesa *grid*, que es la creación de redes.

puede y debe considerarse una infraestructura de estatus comparable o superior al otorgado al *hardware* y redes de computación. Fuente: Griffith, 2005.

Este modelo en capas se ha usado con frecuencia para describir la infraestructura de la información. Originariamente, la infraestructura de la información (a la que algunas veces se hace referencia usando distintos términos) incluye *hardware* computacional, redes y telecomunicaciones, medios de memorias y los *softwares* operativos para que todo pueda trabajar en conjunto. En esta concepción geológica de la información global interconectada y de los recursos de telecomunicaciones, las capas inferiores de los sistemas informáticos y las redes permiten que funcionen las capas más altas de *middleware*, así como las colecciones de contenido digital y aplicaciones. Esta representación gráfica sugiere que el contenido digital, en sí mismo, puede y debe ser considerado una infraestructura de nivel comparable o superior a la que se otorga al *hardware* de cómputo y redes (Griffith, 2005).

La evolución de lo que se denomina *intellectual infrastructures* (Lee, 2008) (en adelante, *infraestructuras intelectuales*, optando por una traducción calcada del inglés) permite entrar a evaluar la incidencia del *copyright* y de los derechos de propiedad intelectual en determinados ámbitos, así como la forma de «construir» excepciones que permitan al mundo desarrollarse más allá de las limitaciones de estas normas. Sobre estos temas se volverá más adelante, en los puntos relacionados con las cuestiones relacionadas con la protección de los datos y, en particular, con la protección resultante de los derechos de propiedad intelectual.

2.1. Las funciones de las ciberinfraestructuras o e-infraestructuras

Una de las primeras cuestiones que debe plantearse es dónde almacenar el conocimiento y de qué manera puede gestionarse; es, por lo tanto, un tema práctico-logístico. No se entrará a analizar cuestiones diferentes de estándares para guardar el conocimiento y poder recuperarlo, ni las cuestiones relacionadas con la protección de los datos de conformidad con el régimen de *copyright* y los derechos de propiedad intelectual, que se analizarán, con detalle, en la parte reservada a la interoperabilidad.

Las formas de almacenar existentes ya no resultan ya ni eficientes ni eficaces para difundir el conocimiento con esta nueva revolución. Se hace necesario crear nuevas infraestructuras, las llamadas infraestructuras del conocimiento o infraestructuras de innovación (Moshe y Teubald, 1986), más operativas para una gestión de datos más eficiente (Lee, 2008).

Estas son infraestructuras que establecen sistemas para gestionar datos a gran escala y a largo plazo con una multiplicidad de usos y de usuarios (Pollock y Williams, 2010). Aparece un nuevo concepto: el de infraestructura en red (Foster, 2002).

Deberá proporcionar mayor funcionalidad que Internet, que es el sistema que le sirve de soporte y, al mismo tiempo, preservar la simplicidad.

Después de un intento de definir lo que son las e-infraestructuras o ciberinfraestructuras, hay que dar un paso más y analizar la función que desempeñan; las más obvias son la producción, el mantenimiento, la distribución, la gestión y la protección del conocimiento (Navarro *et al.*, 2013).

Por lo tanto, una de las funciones de las e-infraestructuras será almacenar, estructurar y hacer accesible el conocimiento. En la medida en que contribuya a hacer accesible el conocimiento y lo ponga a disposición de los consumidores de conocimiento, estos, a su vez, podrán convertirse en generadores de conocimiento. Como ya se ha dicho, las infraestructuras alimentan el crecimiento exponencial del conocimiento con la finalidad de hacer más sencilla la reutilización de los datos y del conocimiento existente, en general.

Por supuesto, las e-infraestructuras, cualquiera que sea el apellido que se les quiera dar, son las garantes del mantenimiento y de la preservación del conocimiento, como lo fueron antes las infraestructuras «físicas», los libros, las bibliotecas, etc. Serán, sin lugar a dudas, las que hagan posible compartir y difundir el conocimiento.

A este respecto, Navarro *et al.* (2013) analizan la siguiente serie de cuestiones:

1. La *fiabilidad* del conocimiento. La multiplicación exponencial de información en la red hace mucho más difícil comprobar todos los datos que pueden servir de base para cada uno de los avances. Es importante, por lo tanto, velar por la fiabilidad de los datos.

Hoy la fiabilidad es clave. Es la medida hasta la cual los resultados son consistentes en el tiempo y una representación veraz de toda la población sujeta a estudio. Si los resultados del estudio pueden reproducirse con una metodología similar, se considera fiable (Joppe, 2000). Incluye, como no puede ser de otra manera, la idea de replicabilidad de los resultados u observaciones. Tal como recoge algún autor (Golafshani, 2003) al citar el trabajo de Kirk y Miller (1986), existen tres tipos de fiabilidad relacionados con la investigación cuantitativa, que hacen referencia a (1) el grado en que se mantiene una medida, dada de forma repetida, (2) la estabilidad de una medida en el tiempo y (3) la similitud de medidas dentro de un determinado periodo de tiempo.

2. La *integridad* del conocimiento. En efecto, es importante identificar las herramientas utilizadas para la construcción de e-infraestructuras e ir trasladando el conocimiento según la tecnología vaya avanzando. De lo contrario, si la infraestructura quedara obsoleta, se perdería todo el conocimiento almacenado

en ella. Sería como perder la llave de acceso a una biblioteca y que, como consecuencia de ello, se perdiera todo el conocimiento almacenado.

La integridad de los datos es cada vez más necesaria; en la medida en que el conocimiento es exponencial, partir de datos erróneos puede afectar a todos los desarrollos subsiguientes. Sin embargo, es difícil de preservar, dada la falta de estandarización en la recogida y almacenamiento de los datos (Mayernik *et al.*, 2008).

3. La *sostenibilidad* del conocimiento a medio o largo plazo. El coste de gestionar y hacer accesible el conocimiento lo están soportando los Estados. En este momento, la sociedad del conocimiento está en pleno proceso de búsqueda de modelos sostenibles que no limiten el acceso a los datos y conocimientos. Aunque es un tema que todavía está muy lejos de haber quedado resuelto, las e-infraestructuras no pueden depender de la financiación pública eternamente.

Hoy, como se verá en distintos ejemplos objeto de esta tesis, lo importante no está solo en la e-infraestructura en sí y en su funcionamiento, sino en su capacidad de interactuar con otras e-infraestructuras. En el fondo, el volumen de datos que se maneja es tan grande que hay que crear metainfraestructuras de datos que puedan interactuar, comunicarse entre sí y aprovechar el conocimiento de unas y otras.

La clave del *big data* radica en permitir que los datos recogidos con una finalidad puedan ser utilizados para una multiplicidad de finalidades distintas. Los datos de salida, los resultados de un sitio, pueden ser los datos de entrada, el principio de otro proyecto; aunque esto siempre ha existido, ahora, tanto la intensidad como la velocidad a la que ocurre es muy superior. Con esa finalidad se están creando las redes *grid*, que no dejan de ser meta e-infraestructuras.

El instrumento jurídico usado hasta ahora eran los tratados internacionales. Sin embargo, la negociación y ratificación de los tratados internacionales hace que la creación de personas jurídicas sea muy complicada.

La aproximación a la creación de e-infraestructuras es diferente en Europa y en Estados Unidos, lo que es un reflejo de las diferentes evoluciones de la regulación en ambos. En efecto, mientras en Europa se ha optado por estructuras de gestión reguladas en normas, en regulaciones, como veremos en el punto siguiente (probablemente como consecuencia de la predominancia de sistemas codificados), en Estados Unidos, en cambio, se ha optado por un sistema más de colaboración y menos regulado entre distintas organizaciones.

Sin embargo, antes del propio análisis de la regulación de las e-infraestructuras, procede analizar de qué manera ha impactado la revolución digital en la ciencia, es decir el nacimiento de la llamada e-ciencia.

2.2. De la ciencia tradicional a la e-ciencia

La aplicación de estas nuevas herramientas digitales a la ciencia da lugar al nacimiento de lo que se conoce como e-ciencia.

En un primer momento, la ciencia se limita a hacer una observación empírica de los fenómenos naturales (Hey *et al.*, 2009). Después, da un paso más, y empieza a teorizar sobre la observación empírica construyendo los primeros modelos y generalizaciones. En las últimas décadas, ha añadido un tercer paradigma: la simulación informática. Hoy se ha dado un cuarto paso: la gestión de datos que unifica teoría, experimentación y simulación (Hey *et al.*, 2009).

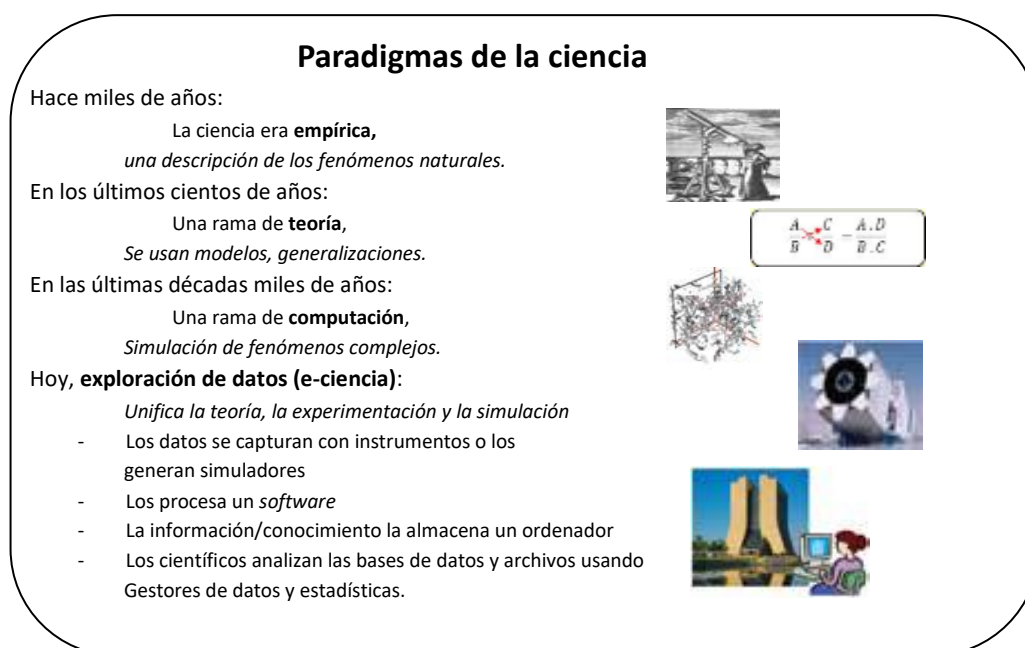


Figura 13. Paradigmas de la ciencia. Fuente: Hey *et al.*, 2009.

El hombre va a tener que enfrentarse a numerosos retos medioambientales en las próximas décadas: el cambio climático, la menor disponibilidad de agua o la pérdida de ecosistemas, por poner algunos ejemplos (Michener, 2011).

La ciencia precisa entender y mitigar esos retos. Para lograrlo va a necesitar un acceso a datos desconocido hasta este momento que le permita cruzar los datos de distintos ámbitos científicos. Los retos principales residen en descubrir datos relevantes, que puedan manejar una enorme heterogeneidad de datos y que puedan convertir los datos en información y conocimiento. Afrontar estos retos requiere nuevos enfoques que permitan gestionar, preservar y analizar. La complejidad de compartir datos es alta y existe tanto desde el punto de vista técnico como desde el punto de vista sociocultural. Uno de los desafíos más importantes es la forma en la que se van a identificar los datos y metadatos dentro y entre los distintos

repositorios así como de qué manera citan los datos tanto los que publican como los que contribuyen con datos.

Es evidente que el paso de la ciencia tradicional a la ciberciencia o e-ciencia tiene el potencial de generar cambios a todos los niveles de la actividad científica, y que estos cambios que se están produciendo en la ciencia son de índole cualitativa (Nentwich y König, 2003).

La e-ciencia está cambiando la forma en la que trabajan los libreros y la forma en que prestan sus servicios (Allard, 2012). Un aspecto importante es, sin duda, que pone el énfasis en el dato.

Una de las funciones importantes de la e-ciencia es la generación de inmensos conjuntos de datos que pueden diseminarse rápidamente a los demás investigadores a través de Internet. En efecto, a día de hoy se han multiplicado los datos recogidos, analizados una y otra vez, almacenados a resultados de los avances tecnológicos, de modelización, etc. Las actividades intensivas en datos han significado un cambio de paradigma. Como consecuencia de ello, surgen una serie de desafíos que necesariamente se tendrá que afrontar y que, posiblemente, van a requerir del trabajo conjunto de científicos, gestores de datos, tecnólogos, etc. En este contexto se está empezando a hablar del cuarto paradigma.

El término ciencia implica el conocimiento adquirido mediante un estudio sistemático; es, por consiguiente, una forma sistemática de construir y organizar el conocimiento en forma de explicaciones verificables. Así, la ciencia de los datos podría centrarse en datos y, por extensión, en estadísticas o en el estudio sistemático de la organización y de las propiedades, y en el análisis de los datos y su papel en la inferencia (Dhar, 2013).

La pregunta que surge, entonces, es por qué es necesario un nuevo término como la ciencia de los datos si, durante siglos, ha existido la estadística. En efecto, el mero hecho de que existan grandes cantidades de datos no es suficiente. Adicionalmente, la materia prima es diferente; ahora se trabajan datos de todo tipo (numéricos, representaciones gráficas, etc.).

Para algunos autores (Pigliucci, 2009) las hipótesis en la ciencia han quedado obsoletas (Anderson, 2008), es decir, que es necesario un enfoque completamente diferente.

La definición de la e-ciencia está en fase de construcción. John Taylor, antiguo director de la Oficina de Ciencia y Tecnología del Reino Unido, decía que la e-ciencia hacía referencia a la colaboración global en áreas clave de la ciencia y que la siguiente generación de infraestructuras hará esto posible (Hey y Trefethen, 2002). Otra definición extraída de esta misma publicación ve la e-ciencia como una síntesis

de información, tecnología y ciencia que permite afrontar retos a escalas antes inimaginables (Yang y Wang, 2011). De lo que no cabe duda es de que las nuevas tecnologías han transformado la investigación (Meyer *et al.*, 2008).

La e-ciencia, a veces, solo hace referencia a las aplicaciones científicas y, otras veces, a las mejoras de todas las disciplinas (Borgman, 2010). En el Reino Unido, la e-ciencia hace referencia a la ciencia a gran escala que está desarrollándose cada vez más a través de colaboraciones que Internet hace posibles. Por lo tanto, una de las características fundamentales de estas empresas colaborativas es que requerirán acceso a grandes cantidades de datos, hacer uso de potentes herramientas de computación y grandes capacidades de visualización para que los usuarios científicos puedan usar tales datos.

En las distintas definiciones de e-ciencia analizadas, se repiten una y otra vez las mismas ideas: colaboración, compartir conocimiento, interconexión y, al mismo tiempo, mayor accesibilidad a los datos, a una mayor cantidad de datos, lo que precisa de herramientas más potentes, más completas.

El proceso de e-ciencia es hoy imparable; hay que encontrar vías para que converja lo que hoy existe con lo que vendrá en el futuro. La ciencia del *big data*, ya hoy, se está convirtiendo en un nuevo motor de la economía (Demchenko *et al.*, 2013) y exige el ejercicio de repensar cómo se construyen las infraestructuras y cuáles son las nuevas soluciones.

Este nuevo enfoque de la gestión de los datos que se mencionaba, así como el manejo de los mismos, que requiere la e-ciencia queda reflejado en el ciclo de vida de gestión de los datos que presentan Demchenko *et al.* (2013).

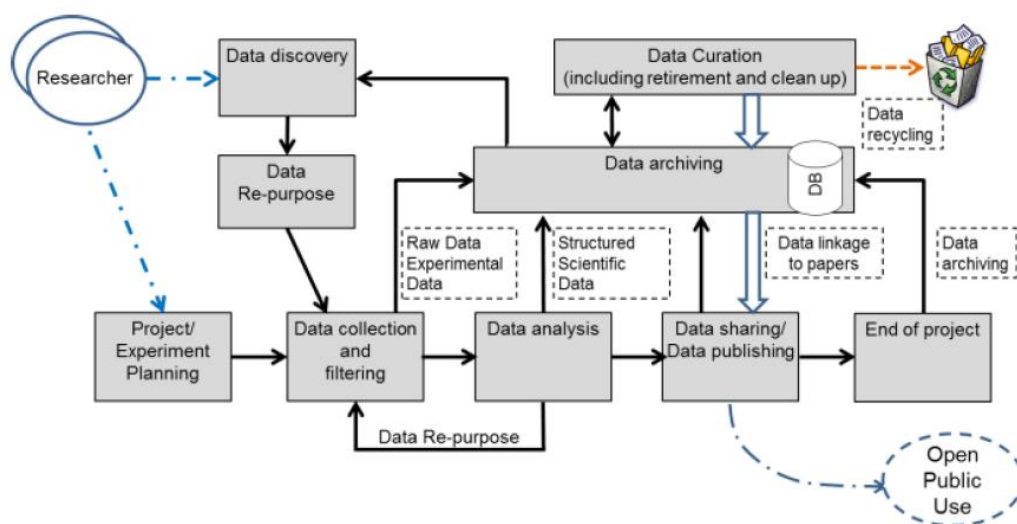


Figura 14. Gestión de los ciclos de vida de los datos científicos en la e-Ciencia. Fuente: Demchenko *et al.*, 2013.

De hecho, analizan una propuesta de ciclo de vida de los datos (Demchenko *et al.*, 2013), aunque señalan que la estructura a la que se tiende irá más al almacenamiento de datos en la nube con un acceso posible desde muchos más puntos. Como se verá más adelante, en el análisis de las e-infraestructuras concretas, la arquitectura está más cerca de la segunda.

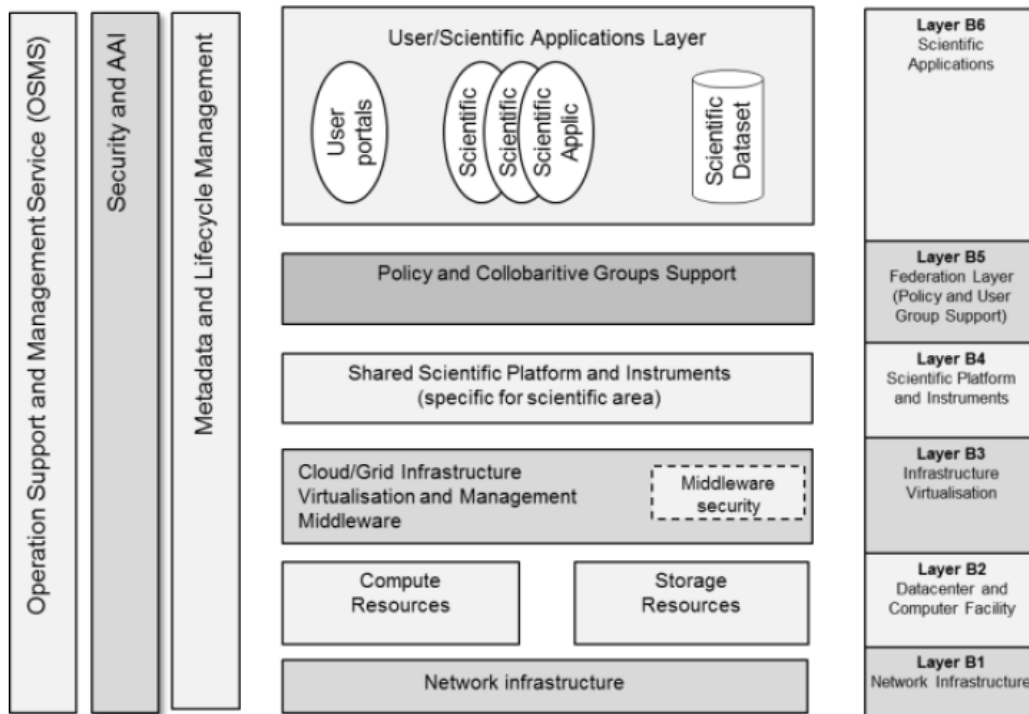


Figura 15. Arquitectura de modelo propuesta. Fuente: Demchenko *et al.*, 2013.

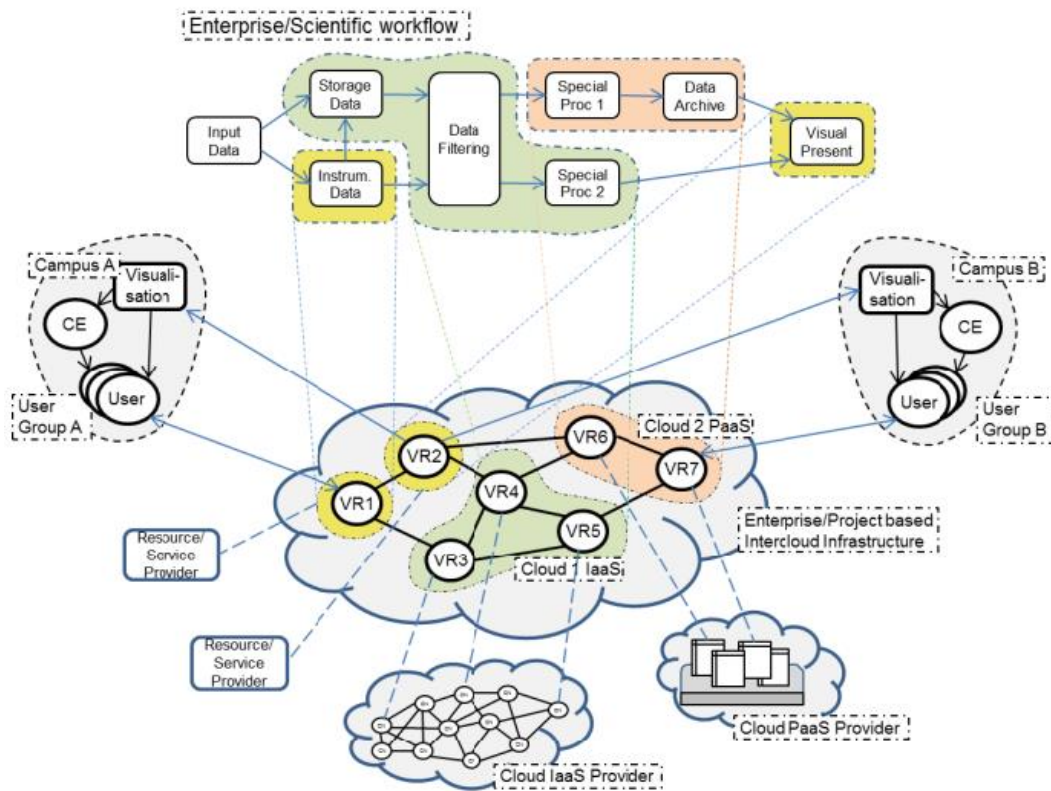


Figura 16. De los flujos de trabajo científicos a la infraestructura basada en la nube. Fuente: Demchenko *et al.*, 2013.

Otra de las cuestiones clave que surge en materia de e-ciencia y de e-investigación es el *open access* o acceso abierto y el hecho de compartir los datos. No obstante, este punto se analizará en profundidad en el punto relativo a la interoperabilidad (Meyer *et al.*, 2008).

Los conceptos de e-ciencia y de e-infraestructura, de alguna manera, se superponen, se confunden. Como consecuencia de esta nueva forma de acercarse al conocimiento, es necesario realizar un análisis de los tipos de problema que se plantean no tanto desde el punto de vista técnico —gestión de los volúmenes de datos, interoperabilidad técnica de las herramientas, la creación de e-infraestructuras, etc., mencionados ya en el punto anterior y sobre los que no se va a profundizar más por no ser objeto de esta tesis—, sino desde el punto de vista jurídico.

Una consideración especial merecerá el concepto de *research data* o dato obtenido en la investigación o dato científico, términos usados con asiduidad por las instituciones españolas y latinoamericanas sobre los que volveremos más adelante, en un epígrafe separado en el apartado de análisis y definición del concepto de dato.

En cualquier caso, debe quedar claro que estos conceptos hacen referencia a datos tanto de las ciencias duras como de las menos duras: las ciencias sociales y las humanidades, así como las ciencias inter y multidisciplinares (Alonso García, 2016).

2.3. Un estudio comparado de dos enfoques distintos en la regulación de las e-infraestructuras

2.3.1. El enfoque en Europa: marco regulatorio, la creación de nuevas formas jurídicas

2.3.1.1. Análisis de los European Strategy Forum on Research Infrastructures (ESFRI)

Como consecuencia de las nuevas necesidades de gestión de datos, aparecieron nuevas estructuras para articular y regular estas nuevas e-infraestructuras con una multiplicidad de denominaciones. Una de estas denominaciones es la de infraestructuras de conocimiento (The Digital Archiving Consultancy Limited, 2008) o infraestructuras de investigación.

En este contexto, aparece el European Strategy Forum on Research Infrastructures (ESFRI). Nació como un instrumento estratégico europeo para el desarrollo de la integración científica en la Unión y el fortalecimiento de su alcance internacional con personalidad jurídica propia. El papel de las estructuras de investigación en el actual progreso de la ciencia y la tecnología es fundamental, así que se constituyó un instrumento clave para reunir a un amplio espectro de agentes implicados en la búsqueda de soluciones a los desafíos de la sociedad de hoy, de modo que se evitaba la necesidad de recurrir a los tratados internacionales.

La mejora y los avances continuados en la salud de los ciudadanos europeos dependen de la investigación biomédica y de la participación ciudadana en esa investigación. La investigación de hoy en día exige recursos nuevos, mejores y de mayor alcance. Más de doscientas instituciones de veinticuatro Estados miembros, incluidas las agencias financiadoras, se han unido para el diseño de e-infraestructuras que proporcionen esos recursos.

Las e-infraestructuras europeas gestionan en la actualidad algo más de diez millones de muestras. El objetivo es aumentar este número de muestras, las bases de datos y las herramientas para incrementar la calidad, reducir la fragmentación y poner todo esto a disposición de todos los investigadores europeos.

A European e-Infrastructure for, and of, e-Science Digital Repositories

Layers of the e-Infrastructure and some of their desired

e-Infrastructure of repositories	Information	Authenticity
	Collections: data, work-flows, publications, learning materials, etc.	Quality
		Longevity
	Repository services	Ease of use
	Deposit, annotation, delivery, visualisation, search, preservation, etc	Availability
		Reliability
	Repositories	Trusted
	Repository management, curation, physical security, etc	Open
		Well managed
	Access	Standardised
	Authentication, authorisation, security, federation, portals etc	Stable
		Flexible
e-Infrastructure for repositories	Management	Transparent
	Grids, virtual organisations, management systems , etc	Responsive
		Informed
	Physical Infrastructure	Available
	Networks, computing, HPC, physical storage, etc	Scaleable
		Reliable

Figura 17. Una e-infraestructura europea para y de repositorios digitales de e-ciencia. Fuente: Comunidades Europeas, 2008.

En diciembre de 2008, a petición del Consejo Europeo de Ministros de Investigación, se actualizó y publicó una nueva hoja de ruta (Comunidades Europeas, 2008). La finalidad de esta hoja de ruta es recoger las actividades de los ESFRI para construir un entorno común, paneuropeo. Es necesario realizar un esfuerzo por parte de todos los Estados miembros y asignar presupuestos nacionales para construir infraestructuras de investigación.

La misión de los ESFRI era servir de pilar para el desarrollo de un enfoque coherente y estratégico a la realización de políticas en lo que respecta a las e-infraestructuras europeas y paneuropeas.

En este contexto, se reveló la inexistencia de instrumentos legales apropiados para la construcción de estas e-infraestructuras de investigación a nivel europeo. Como consecuencia de ello, en las reuniones de la Comisión Europea se discutió un marco legal adecuado. De conformidad con ello, la Comisión presentó una propuesta de reglamento (Consejo de la Unión Europea, 2009).

Además del marco legal, son necesarios instrumentos de financiación: para eso nacieron los distintos programas marco. Los programas marco se crearon como programas de inversión de I+D de la Unión Europea basados en la creación de consorcios para promover la cohesión en Europa. Se pide que concurren distintos países y diferentes tipos de organizaciones: universidades, centros de investigación,

empresas privadas, etc. Sin embargo, existen voces que cuestionan el que se haya conseguido el propósito perseguido de partida (Breschi y Cusmano, 2003).

El último de estos programas marco ha sido Horizonte 2020, que ha conllevado un fuerte esfuerzo de financiación de proyectos realizado conjuntamente por la Comisión Europea y por el Banco Europeo de Inversiones. Además, la Unión Europea está realizando enormes esfuerzos de financiación a través de los fondos estructurales para las nuevas e-infraestructuras de investigación.

De alguna forma, la formalización y publicación de la hoja de ruta de los ESFRI ha contribuido a potenciar la colaboración a nivel internacional, más allá de la Unión Europea. Se publicó por primera vez en 2006, para evitar la duplicación las e-infraestructuras de investigación en la Unión Europea (Calzolari *et al.*, 2014).

El proceso de creación de los ESFRI es de vital importancia en el ámbito de la investigación dentro de la Unión Europea. Las infraestructuras de investigación desarrolladas con la colaboración europea generan oportunidades para que los científicos puedan participar en los intercambios de conocimiento dentro del área de investigación europea (Duşa, 2014). Las e-infraestructuras de investigación son el eje de la ciencia; sin ellas, la ciencia no puede estructurarse.

Este sistema está permitiendo facilitar una mayor interacción con otros países como Australia, Estados Unidos, Japón, Sudáfrica, etc., que se analizarán en el apartado relativo a la bioinformática con ejemplos de proyectos concretos.

The landscape and its projects

Roadmap 2010

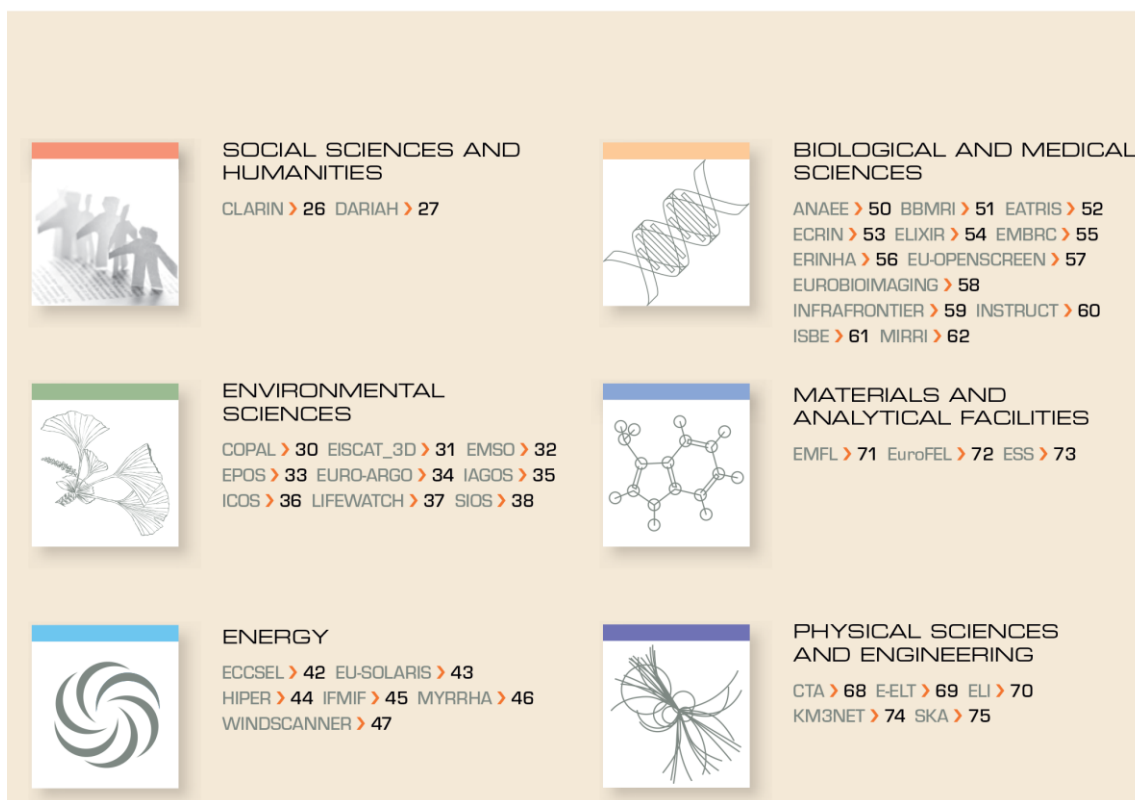


Figura 18. ESFRI Roadmap 2010.

2.3.1.2. Los Consorcios de Infraestructuras de Investigación Europeas

Este concepto se traduce del inglés European Research Infrastructure Consortium (ERIC). Con la creación de estas nuevas e-infraestructuras, en el contexto europeo nace una nueva forma jurídica, precisamente, con la finalidad de facilitar un marco jurídico de referencia: los denominados ERIC. Su creación obedece a la necesidad de armonizar distintos sistemas jurídicos y diferentes culturas, mediante la creación de instituciones supranacionales.

Estos consorcios traen causa de la propia construcción supranacional de la Unión Europea y, por supuesto, del concepto de que las economías de escala son mucho más ventajosas al hacer posible la creación de e-infraestructuras para Estados miembros que no podrían tenerlas de otro modo.

Se va a proceder a analizar de qué manera se constituye un ERIC y su estructura. El Reglamento (CE) n.º 723/2009 del Consejo de 25 de junio de 2009 relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC) crea esta nueva figura jurídica (Consejo de la Unión Europea, 2009).

Los considerandos del reglamento justifican el nacimiento de esta nueva infraestructura. Encuentran la justificación en el artículo 171 del Tratado de la Comunidad Europea que permite a esta «crear empresas comunes o cualquier otra estructura que se considere necesaria para la correcta ejecución de los programas de investigación, de desarrollo tecnológico y de demostración comunitarios». (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 1.º).

El punto de arranque de la creación de los ERIC no es otro que el apoyo a las e-infraestructuras de investigación y a su desarrollo. La Unión Europea se dio cuenta de que el formato de apoyo a través de subvenciones resultaba insuficiente, y buscó implantar un marco jurídico más adecuado para promover y simplificar la creación y explotación de e-infraestructuras de investigación.

Así, al amparo del artículo 171 del tratado, fueron creados los ERIC que el considerando sexto del reglamento define como: «una entidad jurídica de la que la Comunidad no forma parte necesariamente y a la que esta no aporta una contribución financiera, [...]». No obstante, prosigue el reglamento en su considerando séptimo:

[D]ebe corresponder a los Estados miembros interesados, solos o en colaboración con otras entidades habilitadas, definir sus necesidades en materia de creación de e-infraestructuras de investigación en esta forma jurídica, sobre la base de sus actividades de investigación y desarrollo tecnológico y de los requisitos de la Comunidad. Por esta misma razón, la afiliación a un ERIC debe estar abierta a los Estados miembros interesados, con la posible participación de países habilitados asociados al programa marco de investigación y desarrollo tecnológico [...] y de terceros países y de organizaciones intergubernamentales especializadas. Además de miembros de pleno derecho, los Estados miembros deben poder constituirse en observadores de un ERIC en las condiciones especificadas en los Estatutos.

La función principal de un ERIC es la «creación y explotación de infraestructuras de investigación sin fines lucrativos y debe dedicar la parte fundamental de sus recursos [...]» a ello (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 8.º).

O como se señala en Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009 (cdo. 9.º):

Las infraestructuras de investigación deben contribuir a salvaguardar la excelencia científica de la investigación de la Comunidad y su competitividad económica sobre la base de previsiones a medio y corto plazo, a través de un apoyo eficiente a las actividades europeas de investigación. A tal fin, deben

estar legalmente abiertas a la comunidad científica europea en general de acuerdo con las normas fijadas por sus estatutos, tener por objetivo potenciar la capacidad científica europea más allá de la situación actual y contribuir, por lo tanto, al desarrollo del espacio europeo de investigación.

Por motivos de transparencia, el ERIC debe crearse mediante una decisión y publicarse en el Diario Oficial de las Comunidades Europeas (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, 11.º). Debe dotarse al ERIC de personalidad jurídica y de la capacidad jurídica más amplia posible, para que pueda trabajar en la consecución de sus fines. (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 21.º).

Cada ERIC debe contar, al menos, con tres Estados miembros y puede incluir países asociados y terceros países no asociados u organizaciones intergubernamentales especializadas (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 13.º). Teniendo en cuenta la dimensión comunitaria de la norma, los Estados miembros que formen parte del ERIC deben tener mayoría de votos en la junta de afiliados (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 14.º). Los ERIC deben disponer de la flexibilidad necesaria para modificar sus estatutos sin que la Comunidad, que es la que crea el ERIC, pierda el control de los elementos esenciales. Cualquier modificación esencial del ERIC deberá aprobarla la Comunidad antes de que surta efectos (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 15.º).

Los ERIC deben dotarse de órganos propios, a los efectos de realizar una gestión efectiva de sus actividades. Los estatutos deben determinar la forma en la que dichos órganos representan legalmente al ERIC (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 16.º).

Los ERIC deben llevar a cabo sus actividades con principios financieros saneados y tendrán responsabilidad financiera (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 17.º).

Será aplicable a los ERIC la ley del Estado en el que tengan su sede estatutaria, y el derecho comunitario, que es de donde procede su norma de creación (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 21.º).

Los ERIC deben rendir cuentas a la Comisión; en caso de que no lo hagan la Comisión puede decidir su liquidación (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, cdo. 23.º).

El artículo 4 del Reglamento (CE) n.º 723/2009 del Consejo de 25 de junio de 2009 relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras

de Investigación Europeas (ERIC) establece los requisitos relativos a la e-infraestructura:

La infraestructura de investigación que vaya a establecer el ERIC deberá cumplir con los requisitos siguientes:

- a la difusión y utilización óptima de los resultados de las actividades comunitarias de investigación, de desarrollo tecnológico y de demostración.

Los ERIC deberán elaborar unos Estatutos de conformidad con el artículo 10 del Reglamento (CE) n.º 723/2009 del Consejo de 25 de junio de 2009 relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC). Estos estatutos, como mínimo, deberán recoger los siguientes puntos.

1. La lista de afiliados, observadores y, en su caso, de las entidades que representan a los miembros y las condiciones y procedimiento aplicables a la modificación de la lista de afiliados del ERIC.
2. El cometido y las actividades del ERIC.
3. La sede estatutaria (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, art. 8.1).
4. La denominación del ERIC (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, art. 8.2).
5. La duración del ERIC y el procedimiento de liquidación (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, art. 16).
6. El régimen de responsabilidad (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, art. 14).
7. Los principios básicos que regulen:
 - la política de acceso de los usuarios,
 - la política de evaluación científica,
 - la política de difusión,
 - la política sobre derechos de propiedad intelectual,
 - la política de empleo, incluida la igualdad de oportunidades,
 - la política de contratación pública, que debe respetar los principios de transparencia, no discriminación y competencia,
 - en su caso, el desmantelamiento de las instalaciones,

- la política en materia de datos.
- 8. Los derechos y obligaciones de los afiliados, incluidos la obligación de aportar contribuciones a un presupuesto equilibrado y los derechos de voto.
- 9. Los órganos del ERIC, sus funciones, responsabilidades, el modo en que están constituidos y los procedimientos de decisión, especialmente en lo que atañe a la modificación de los estatutos, de conformidad con los artículos 11 y 12.
- 10. La lengua o lenguas de trabajo.
- 11. Las referencias a las disposiciones de aplicación de los estatutos.

Además, cada ERIC deberá elaborar un informe anual de actividades donde expondrá los aspectos científicos, operativos y financieros (Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009, art. 17).

El Reglamento regulador de los ERIC ha sufrido una modificación en el Reglamento (UE) n.º 1261/2013 del Consejo de 2 de diciembre de 2013, por el que se modifica el Reglamento (CE) n.º 723/2009 relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC) en lo que respecta al artículo 9 (Consejo de la Unión Europea, 2013).

El número de ERIC en Europa es limitado y estos son los que existen en la actualidad, según la Comisión Europea (4 de diciembre de 2017):

- European Multidisciplinary Seafloor and water-column Observatory (EMSO), cuya función es la monitorización a largo plazo de los procesos medioambientales relacionados con la interacción entre la geosfera, la biosfera y la hidrosfera, incluyendo los riesgos naturales.⁶
- Integrated Carbon Observation System European Research Infrastructure Consortium (ICOS).⁷
- European Spallation Source-ERIC (ESS).⁸
- Joint Institute for VLBI (JIV). Un nuevo consorcio de megantenass que llevará la astronomía a un nuevo nivel.⁹

⁶ Para más detalle, puede consultar la página web del organismo disponible en <http://www.emso-eu.org>. (Consultado por última vez el 1 de abril de 2017).

⁷ Para más detalle, puede consultarse la página web <https://www.icos-ri.eu>. (Consultado por última vez el 1 de abril de 2017).

⁸ Para más información, puede consultarse la página web <https://europeanspallationsource.se>. (Consultado por última vez el 1 de abril de 2017).

- DARIAH, que es una infraestructura europea digital para que los estudiosos de artes y humanidades puedan trabajar con métodos computacionales. Soporta la investigación digital y la enseñanza con medios digitales.¹⁰
- Central European Research Infrastructure (CERIC): un intento de unificar las instalaciones científicas de Centroeuropa con la finalidad de ofrecer un servicio integrado a los investigadores.¹¹
- European contribution to Argo program (Euro ARGO).¹²
- European Clinical Research Infrastructure Network (ECRIN-ERIC).¹³ ECRIN es una organización pública, sin ánimo de lucro, que relaciona socios científicos dentro de Europa con la finalidad de facilitar la investigación clínica multinacional. Asesora y proporciona servicios de gestión y herramientas a patrocinadores e investigadores para que puedan superar las trabas que surgen en los ensayos multinacionales y promover la colaboración.
- Biobanking and Biomolecular Resources Research Infrastructure (BBMRI).¹⁴ Pone en común los biobancos y los recursos biomoleculares en una infraestructura paneuropea.
- European Social Survey (ESS),¹⁵ que persigue explorar las actitudes sociales, medir los cambios en las actitudes en público y analizar los patrones de comportamiento.
- European infrastructure for translational medicine (EATRIS).¹⁶ Esta infraestructura persigue lograr que se traduzcan los descubrimientos científicos en productos médicos más efectivos.

⁹ Si se quiere ampliar la información se puede consultar <http://www.jive.nl>. (Consultado por última vez el 1 de abril de 2017).

¹⁰ Puede consultarse <http://www.dariah.eu>. (Consultado por última vez el 1 de abril de 2017).

¹¹ En la página web <http://www.ceric-eric.eu> puede ampliarse la información aquí citada. (Consultado por última vez el 1 de abril de 2017).

¹² Para más detalle, puede consultarse la página <http://www.euro-argo.eu>. (Consultado por última vez el 1 de abril de 2017).

¹³ Para tener una información más detallada sobre el tema, consúltese <http://www.ecrin.org>. (Consultado por última vez el 1 de abril de 2017).

¹⁴ Para más detalle, véase <http://www.bbmri-eric.eu>. (Consultado por última vez el 1 de abril de 2017).

¹⁵ Para más información, puede consultarse la página web <http://www.europeansocialsurvey.org>. (Consultado por última vez el 1 de abril de 2017).

- European Research Infrastructure for Language Resources and Technology (CLARIN).¹⁷ Esta infraestructura pone los recursos digitales de lenguaje de todas las disciplinas, en particular de las disciplinas relacionadas con las humanidades y las ciencias sociales, a disposición de los estudiosos, investigadores, estudiantes y de los ciudadanos científicos mediante un acceso único. Ofrece soluciones a largo plazo y servicios tecnológicos para desplegar, conectar, analizar y dar apoyo a los datos y herramientas de lenguaje digital. Persigue contribuir a la construcción de una verdadera área europea de investigación multilingüe.
- Survey of Health, Ageing and Retirement in Europe (SHARE).¹⁸ Es una base de datos multidisciplinar y multinacional de datos de salud, de estatus socioeconómico y de redes familiares y sociales de aproximadamente 123 000 individuos de cincuenta años o más (más de 293 000 entrevistas. Cubre veintisiete países europeos e Israel.
- European Infrastructure for Chemical Biology (EU-OPENSOURCE).¹⁹ Es una infraestructura que da apoyo a la investigación de las ciencias de la vida y a cómo estas se traducen a la medicina y a la agricultura.
- Council of European Social Science Data Archives (CESSDA).²⁰ Proporciona servicios de datos integrados y sostenibles a gran escala de las ciencias sociales. Pone en relación los archivos de ciencias sociales dentro de Europa con la finalidad de promover los resultados de la investigación de ciencias sociales y para dar soporte a la investigación nacional e internacional y a la cooperación.
- European Carbon Dioxide Capture and Storage Laboratory Infrastructure(ECSEL).²¹ Su visión es reducir las emisiones de la industria de dióxido de carbono y contribuir a combatir el cambio climático global.

¹⁶ En la página web <http://eatris.eu> puede ampliarse esta información. (Consultado por última vez el 1 de abril de 2017).

¹⁷ Para más detalle, consultar <https://www.clarin.eu>. (Consultado por última vez el 1 de abril de 2017).

¹⁸ Si se quiere ampliar la información se puede consultar <http://www.share-project.org>. (Consultado por última vez el 1 de abril de 2017).

¹⁹ Véase, si se quiere ampliar esta información la siguiente página web <http://www.eu-openscreen.eu>. (Consultado por última vez el 1 de abril de 2017).

²⁰ Para más detalle, consúltase <https://cessda.net>. (Consultado por última vez el 1 de abril de 2017).

²¹ si se quiere ampliar esta información, véase <http://www.eccsel.org>. (Consultado por última vez el 1 de abril de 2017).

- Science and Technology Infrastructure for Research on Biodiversity and Ecosystems (LifeWatch).²² Esta última infraestructura será objeto de un análisis más detallado en el punto relacionado con la bioinformática.

2.3.2. El enfoque en Estados Unidos: las políticas estadounidenses

A diferencia del régimen que existe en Europa, en Estados Unidos no hay una regulación de carácter general, sino un análisis caso por caso a través de contratos *ad hoc* con la Administración Pública, y la concesión de subvenciones concretas aplicables a cada caso de forma individual. La regulación se hace mediante lo que los estadounidenses denominan *grants*, es decir, subvenciones (Allard, 2012).

Los productores de datos científicos pueden agruparse en tres grupos (Uhlir y Reichman, 2003):

- Las agencias gubernamentales.
- Las organizaciones académicas y otras organizaciones de investigación sin ánimo de lucro.
- El sector privado con fines comerciales.

En Estados Unidos, a diferencia de lo que ocurre en otros lugares (por ejemplo: en Europa), en relación con el sistema de protección (*copyright* o derechos de propiedad intelectual e industrial), existe una importante atención en el dominio público. Eso favorece que los hallazgos realizados tanto por las agencias gubernamentales como por otras organizaciones sin ánimo de lucro se pongan en el dominio público o, cuando menos, sean más fácilmente accesibles (*fair use*).

De alguna manera, el acceso a las llamadas *grants* supone aceptación de que los resultados, los hallazgos, los datos que formen parte de la investigación pasen a formar parte del dominio público, de acuerdo con las llamadas políticas de *public access* muy desarrolladas, que se analizarán en profundidad en el punto relativo a la interoperabilidad legal.

Asegurarse de que los datos ligados a la investigación son accesibles, para que puedan usarse el mayor número de veces y de la forma más amplia posible es un tema de gestión racional de recursos públicos (Arzberger *et al.*, 2004).

No obstante, no todos los datos se ponen en el dominio público —sobre el concepto de dominio público se volverá en un punto concreto más adelante; ahora, procede analizar a que datos afecta—. A estos efectos, diferencian tres tipos de

²² Para ampliar esta información, véase <http://www.servicecentrelifewatch.eu/home>. (Consultado por última vez el 1 de abril de 2017).

datos, en función de su origen. Cada uno recibirá un tratamiento diferenciado (Uhlir y Reichman, 2003).

3. LAS CUESTIONES JURÍDICAS QUE PLANTEA LA REVOLUCIÓN DIGITAL

El análisis del primer marco contextual conduce a poner encima de la mesa todas las cuestiones jurídicas que esta revolución digital plantea y lo que ello significa. La revolución digital, no solo trae nuevas instituciones para la gestión de las nuevas situaciones, sino que afecta a las preexistentes con un impacto en su operatividad que ha requerido una serie de cambios que es importante analizar.

Existe una resonancia entre las necesidades de la ciencia de la biodiversidad y las oportunidades de globalización y de interoperabilidad que proporciona Internet (Bisby, 2000). Por ello, en este apartado se van a ir analizando las distintas cuestiones que surgen, se contextualizarán las problemáticas para terminar con un análisis de la construcción del concepto de interoperabilidad general, primero, y legal, después.

3.1. Una primera reflexión sobre las cuestiones jurídicas previas a la interoperabilidad

3.1.1. El concepto de dato: su estatuto jurídico

Surge toda una serie de problemáticas en relación con los datos y las transacciones que se producen con ellos. El análisis de la definición de dato es la condición previa antes de entrar a valorar los problemas que surgen en el marco de la interoperabilidad. En efecto, el dato es casi el paso previo, lo que se introduce en la red. Es una representación simbólica (numérica, alfabética, algorítmica, etc.) de un atributo o variable cuantitativa. Los datos describen hechos empíricos, sucesos y entidades (Dato, s. f.).

Acotar el término «dato» resulta relevante puesto que, en función de lo que incluya la definición, el impacto de los derechos de propiedad intelectual será diferente. Sin embargo, su definición está cada vez menos clara (Fishbein, 1991).

Existe una diferenciación básica entre dato, información y conocimiento. La visión generalmente aceptada es la que considera los datos como simples hechos que se convierten en información a medida que los datos se combinan en estructuras significativas, que posteriormente se convierten en conocimiento cuando la información significativa se pone en un contexto y cuando se puede usar para hacer predicciones. Esta visión ve los datos como un requisito previo para la información, y la información como requisito previo para el conocimiento (Tuomi, 1999).

Por ejemplo, para algunos autores, los datos tienen la consideración de símbolos que aún no han sido interpretados; la información es un dato con significado y el conocimiento es lo que permite a las personas asignar significado y, de ese modo, generar información (Spek y Spijkervet, 1997). Otros consideran que los datos son observaciones simples de estados del mundo, la información son los datos dotados de relevancia y propósito y el conocimiento es información valiosa (Davenport, 1997).

De forma intuitiva, está claro que para construir conocimiento hay que partir de la observación de hechos simples (Tuomi, 1999). Un concepto muy controvertido es el de los «datos brutos»; algunos autores defienden que no existen en la medida en que en la propia captura de los datos ya existe una percepción influenciada por usos potenciales, expectativas, contexto y construcciones teóricas (Tuomi, 1999).

A toda esta discusión hay que añadir otra abierta entre los científicos a día de hoy: si el término «dato» hace referencia solo a los resultados de la investigación o también a los datos recogidos durante el proceso de elaboración de la observación.

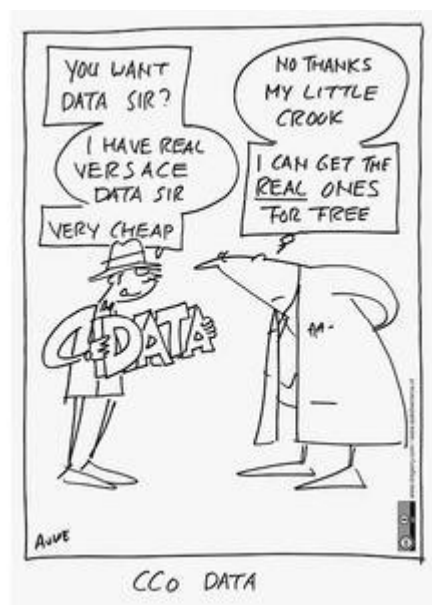


Figura 19. Dibujo de Auke Herrema²³.

Resulta también controvertido el tema de si, además de los resultados, deben publicarse los datos que han permitido llegar a los mismos, de manera que pueda profundizarse la investigación o incluso cuestionarse, sin necesidad de volver a arrancar de cero.

²³ Disponible en: www.fosteropenscience.eu/content/cartoon-cco-data. (Consultado por última vez el 5 de marzo de 2018).

Los datos en la ciencia son como ladrillos, y los conceptos teóricos son el mortero que los conecta para estructurarlos. Cada nuevo bit de datos es una pieza: puede utilizarse para demostrar una hipótesis, elaborada a partir de información, o recogida en observaciones, experimentos o cálculos. Como una pieza observada o medida, se convierte en parte de nuestra base de conocimiento, para compartir, interpretar y conciliar con los datos ya existentes. Los científicos se preguntan:

¿Son estos nuevos datos consistentes con lo que ya sabemos? ¿Son justo lo que podía esperarse, o nos obligan a cuestionar los resultados, a repetir el experimento, o encontrar una nueva interpretación que explique por qué Los datos son lo que son? (Committee on Issues in the Transborder Flow of Scientific Data, 1997).

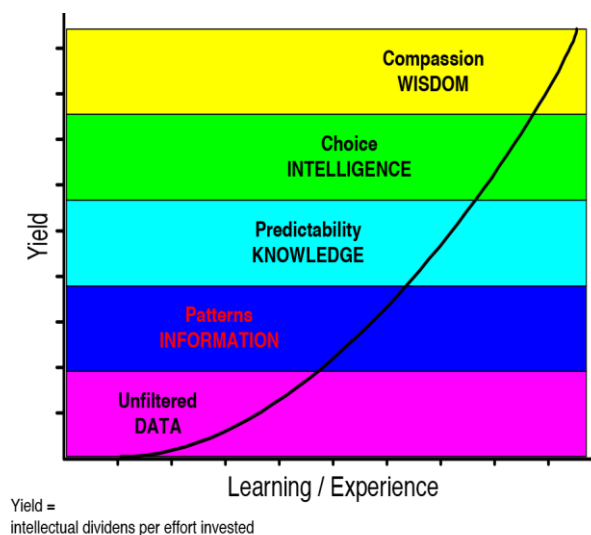


Figura 20. Visión convencional de la jerarquía del conocimiento. Fuente: Tuomi, 1999.

3.1.2. El concepto de dato científico (research data)

El concepto de dato científico, cuya definición plantea dificultades, se ha importado del término inglés *research data*, que equivaldría, en realidad, a «datos en proceso de investigación». No obstante, para simplificar la cita, en adelante se hará referencia a «dato científico. Los datos en proceso de investigación o datos científicos son los datos recogidos, observados o creados con la finalidad de analizar y generar resultados originales (Boston Library, s.f.).

El concepto de dato científico incluye una multiplicidad de formatos como pueden ser los textos de Word, las tablas de Excel, las notas de laboratorios, los datos recogidos en el campo, los diarios, cuestionarios y transcripciones, las grabaciones de video y de audio, fotografías y películas, las repuestas a cuestionarios, las transparencias, los especímenes y muestras, las colecciones de objetos digitales adquiridos o generados durante el proceso de investigación, los archivos de datos, los contenidos de las bases de datos incluidos los videos, audios, textos e imágenes,

los modelos y algoritmos, los contenidos de una aplicación como pueden ser los *inputs*, *outputs*, *log files* de los análisis de *software*, los *softwares* de simulación, las metodologías y flujos, los procedimientos estándar de operación y los protocolos, etc. No es una definición exhaustiva, pero sí una forma de ilustrar qué tipos de datos pueden tener esta consideración.

En el estudio, los datos científicos quedan definidos como una colección de datos digitales estructurados de cualesquiera disciplinas a fuentes de los que los investigadores pueden hacer uso en la realización de sus investigaciones o que les pueden servir para llevar a cabo su investigación (Beagrie *et al.*, 2009). Su origen puede ser muy diverso: pueden proceder de fuentes totalmente al margen del ámbito académico o de fuentes relacionadas con este. Incluyen una gran variedad y heterogeneidad de datos y sus correspondientes metadatos, documentación para hacerla utilizable y comprendida, o las representaciones y registros digitales para datos científicos. En esencia, cualquier tipo de datos científicos almacenados en los repositorios de datos estarían en su alcance.

Los ejemplos podrían incluir: datos complejos utilizados en modelos climáticos, aerodinámica, modelado molecular, bioinformática, archivos de video e imágenes utilizados en arqueología, historia del arte, antropología, imágenes digitales o datos de investigación de fuentes físicas primarias en humanidades, datos cuantitativos y cualitativos utilizados en las ciencias sociales o datos e índices electrónicos para fósiles o muestras de tejido cutáneo (Beagrie *et al.*, 2009).

Dentro del espectro de los datos en proceso de investigación existe una serie de categorías que procede diferenciar. Cada una de las categorías puede necesitar un distinto tipo de plan de gestión de datos.

1. Datos derivados de la *observación*, es decir, datos capturados en tiempo real, por lo general, insustituibles. Por ejemplo: los datos obtenidos de los sensores, los datos topográficos, los datos de muestra o las imágenes neurológicas. Reciben también el nombre de datos en bruto o inicialmente procesados, por ejemplo, en un centro de investigación como puede ser un observatorio.
 - Por lo general, se ponen a disposición del público en este formato después de un periodo de embargo de por ejemplo un año.
 - En algunos casos, pueden ponerse a disposición del público inmediatamente, por ejemplo: Swift Gamma Ray Burst.
2. Datos derivados la *experimentación*, es decir, datos obtenidos de equipos de laboratorio que, con frecuencia, pueden reproducirse, pero cuya reproducción puede ser muy costosa. Por ejemplo: secuencias de genes,

cromatógrafos o datos de campos magnéticos. Son los datos científicos preparados para la investigación, totalmente calibrados, combinados y anotados.

- A menudo se producen por individuos o mediante colaboraciones.
 - Rara vez están disponibles para cualquier persona fuera de la colaboración, excepto a petición o mediante colaboración.
 - Es necesario, si se desea reutilizar para la ciencia, a menos que se tengan conocimientos específicos de subdominios e información contextual detallada para reproducirse a partir de datos en bruto.
 - Se consideran una forma de hacer posible una ventaja competitiva para los investigadores involucrados.
 - Pueden servir para generar futuras muestras y documentos adicionales para la colaboración, además, de los resultados originales publicados.
 - En algunos casos, pueden ser producidos por científicos de datos dedicados en nombre de la comunidad.
3. Datos derivados de la *simulación* son los datos obtenidos de modelos de pruebas en los que modelos y metadatos son más importantes que los datos de salida. Por ejemplo: modelos de clima o modelos económicos. Son los conjuntos de datos de salida publicados tras un detallado análisis de conjuntos de datos científicos preparados para la investigación.
- Con datos bajo el gráfico en una publicación de la revista después del análisis de conjuntos de datos científicos.
 - Pocas veces están disponibles para cualquier persona al margen de la colaboración excepto a petición.
 - Es muy posible que generen muestras futuras adicionales y papeles para la colaboración además de los originales.
 - Puede ocurrir que otros investigadores soliciten los datos para su propia investigación y no los obtengan.
4. Tipo de catálogo publicado con la representación del conjunto de datos de salida publicados.
- Datos *derivados* o *compilados*: los datos pueden reproducirse, pero el proceso es caro. Por ejemplo: el *text and data mining* (que se

analizará después con más detalle), las compilaciones de bases de datos o los modelos en 3D.

- Datos de *referencia* o *canónicos*: una conglomeración (estadística y orgánica) o una colección de pequeños conjuntos de datos (revisados por pares), probablemente publicados y preservados. Por ejemplo: la secuenciación de genes, los bancos de datos, las estructuras químicas o portales de datos espaciales.

A continuación, se va a proceder a analizar dos enfoques diferenciados en lo que respecta a la regulación de las e-infraestructuras: el europeo y el estadounidense. Sin embargo, sigue sin existir consenso a la hora de definir lo que es un dato científico (*research data*) (Alonso García, 2016).

El núcleo duro del concepto es claro, prosigue Alonso García (2016): es un dato obtenido de la realidad física arrancando del esfuerzo realizado para descubrirlo con metodología científica (y, en su caso, inversión tecnológica de apoyo). Todo dato o serie de datos creada, cuidada (*curated*) y transmisible electrónicamente por la web (o incluso en soporte no digital) puede convertirse y es, rutinariamente también, objeto de investigación científica; desde los datos del censo de población y el resto de la estadística pública hasta los puros datos sobre los que se ha basado la investigación si son susceptibles de protegerse tales como los que fundamentan la solidez de estudios o publicaciones, informes etc., encargados por los distintos poderes públicos o por ellos financiados (salvo cláusula en contrario que no debe ser de adopción discrecional, sino estar basada en que su no adquisición abarata la inversión pública siendo urgente la obtención de la prestación y hacerlo a coste razonable). Va ligado a la divulgación en abierto de estos datos, aunque no sean calificables de «datos científicos» (pero sin que nadie pueda impedir que se conviertan en tales).

Así, por ejemplo, los datos de investigación son aquellos datos que se usan como fuentes primarias para apoyar la investigación técnica o científica, las actividades de investigación de índole académica o artística, y como evidencias en el proceso de investigación o que son comúnmente aceptados por la comunidad de investigación para validar los resultados de la investigación. Todos los contenidos digitales y no digitales tienen el potencial de convertirse en datos de investigación. Los datos de investigación pueden ser datos experimentales, datos de observación, datos operacionales, datos de terceros, datos del sector público, datos de seguimiento, datos procesados o datos reutilizados.²⁴

²⁴ Disponible en http://digital.csic.es/bitstream/10261/138273/1/Datos2016_Bernal.pdf. (Consultado por última vez el 14 de marzo de 2018).

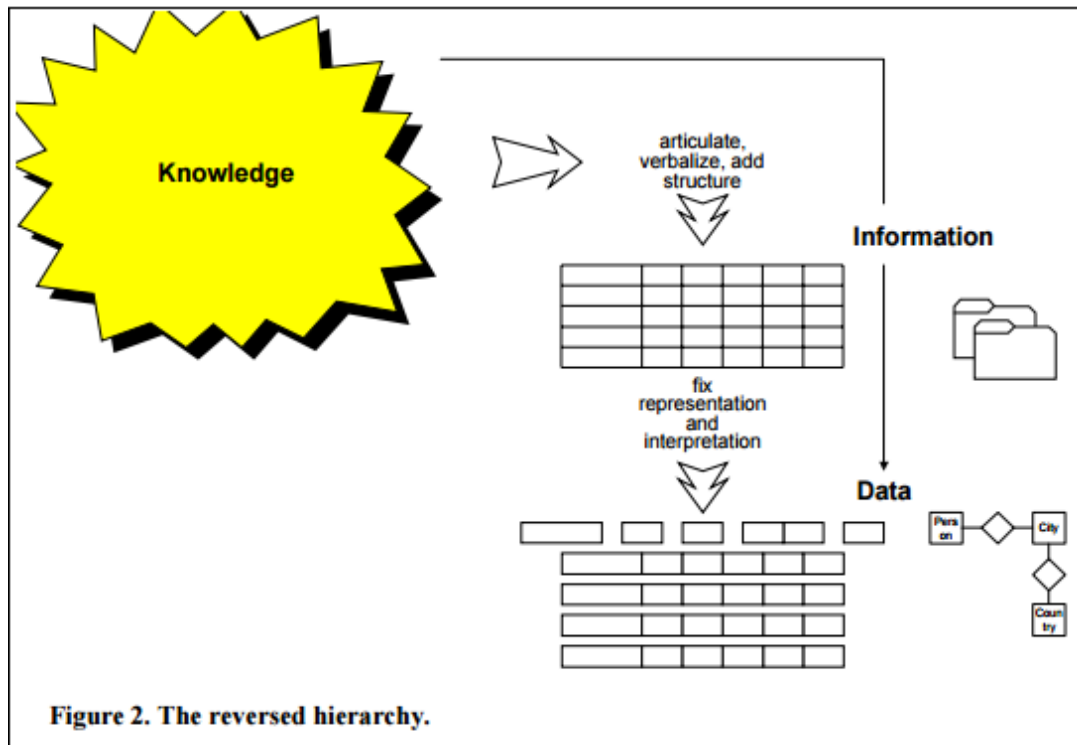


Figura 21. La jerarquía inversa. Fuente: Tuomi, 1999.

Los datos, por lo tanto, no se convierten en información una vez que se les da significado. Por el contrario, tienen su origen en la información, es decir, mediante la introducción de información en una estructura de datos que define completamente su sentido. En lugar de materia prima para la información, los datos resultan de agregar valor a la información poniéndola de forma que se puedan procesar automáticamente (Tuomi, 1999).

Por lo tanto, el proceso se convierte en un ciclo. Así, el aluvión de datos existentes hoy en día en lo que respecta a la investigación (de biodiversidad, objeto de esta tesis) está haciendo que el dato y su gestión sean conceptos distintos (Conner *et al.*, 2013). El método tradicional de gestión de los datos se muestra en la figura siguiente.

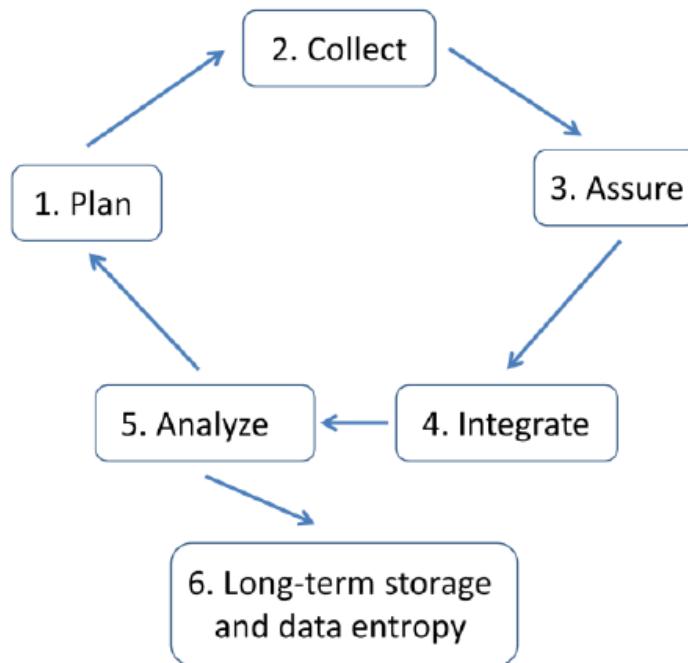


Figura 22. Ciclo de datos tradicional. Fuente: Conner *et al.*

Mientras que en la actualidad, la gestión de los datos es algo diferente, como se representa en la figura 22.

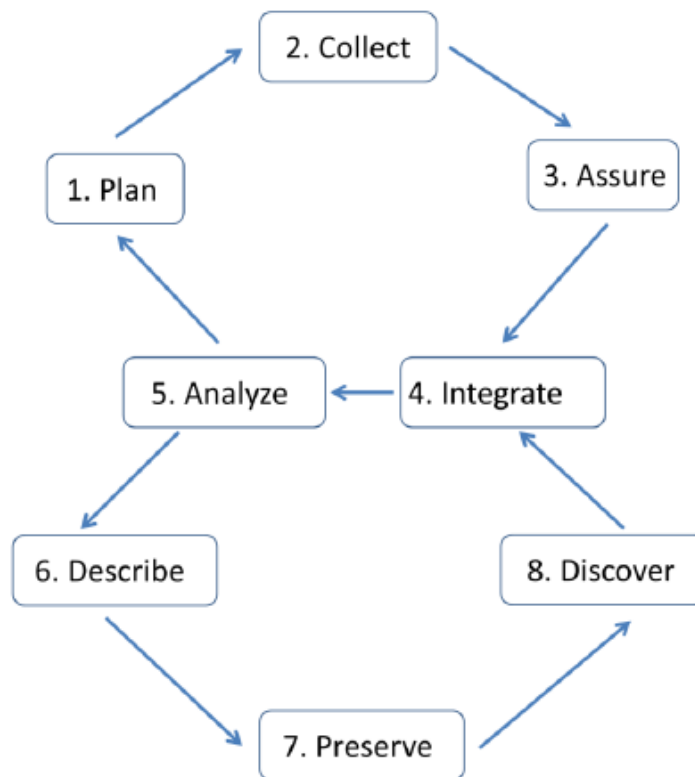


Figura 23. Ciclo de vida de los datos extendido. Fuente: Conner *et al.*

La problemática de los datos relacionados con la diversidad va un paso más allá. En efecto, de alguna manera, los datos de biodiversidad son observaciones de la naturaleza. De hecho, algunos autores como Costello *et al.* (2013) afirman que los datos de biodiversidad tienen que estar disponibles para todos salvo que existan razones de peso para que no lo estén. Incluso, algunos autores como Kelling *et al.* (2009) consideran que la clave del conocimiento reside hoy en los datos en sí mismos.

Más allá de los meros datos, se habla cada vez más de los metadatos. Este término hace referencia a los datos, pero también a las estructuras de datos. Existen definiciones diferentes en función del estadio de gestión de conocimiento del que se esté hablando. Si los datos son el elemento básico, los metadatos constituyen integraciones de datos que acercan a la gestión del conocimiento. Shoit Ikematu (2001) reconoce que existen distintas interpretaciones y recoge como claves las siguientes:

- Los metadatos son datos que describen los atributos de un recurso. Soportan un número de funciones: localización, descubrimiento, documentación, evaluación, selección, etc.
- Los metadatos proporcionan un contexto para entender los datos a través del tiempo.
- Los metadatos se dan asociados con objetos que ayudan a sus potenciales usuarios a obtener ventajas del conocimiento de su existencia o características.
- Los metadatos son el instrumento para transformar los datos brutos en conocimiento.

Por lo tanto, los metadatos son los datos elaborados, preparados y, muy posiblemente, pueden considerarse base de la interoperabilidad.

3.1.3. Los obstáculos jurídicos en relación con el concepto de dato

3.1.3.1. La propiedad de los datos

Después de una aproximación a la definición de dato y a las cuestiones que de la misma se derivan, el siguiente tema que, con carácter previo, merece una reflexión es el de la propiedad de los datos. Existe un gran debate doctrinal en relación con la propiedad de los datos (Culliton, 1988) que acaba cristalizando en dos grandes corrientes. Una de las corrientes, la más clásica, se aferra a la propiedad del dato

ligada al soporte en el que el dato se encuentra y defiende que existe una propiedad de los datos; diferencia dos posibles «propietarios». La corriente más tradicional considera que la propiedad es de los científicos, de los investigadores. Aunque, en un primer momento, puede parecer adecuado que los datos sean propiedad de las personas que han realizado la investigación, otra parte de la doctrina considera que la propiedad, sin lugar a dudas, solo puede ser de las entidades que financian los proyectos (Fishbein, 1991).

Sin embargo, aunque eso pueda ser cierto históricamente, los «propietarios» — Estados, universidades, etc.— no han reclamado la titularidad de esos datos, por lo que son los científicos los que deciden lo que se publica. Así, puede parecer que existe una justificación para el interés de los científicos en «apropiarse de los datos» (Jones, 1986). Al fin y al cabo, es un requisito para proteger las publicaciones que quieran hacer. Sin embargo, muchos autores discrepan.



Figura 24. Dibujo de Auke Herrema²⁵.

En cualquier caso, lo relevante a la hora de analizar la susceptibilidad de apropiación de los datos sería analizar el origen de los mismos. No es lo mismo hablar de datos obtenidos en el marco de proyectos financiados con fondos públicos que de datos obtenidos en el marco de proyectos del sector privado. Como consecuencia de ello, su tratamiento será diferente.

Es evidente que es difícilmente discutible el que la empresa privada que invierta en proyectos pueda explotar los datos obtenidos en el mismo. Otra cosa serán los datos vinculados a investigaciones subvencionadas con fondos públicos. Cada vez cobra

²⁵ Recuperado de <https://bit.ly/2HJMooc>. (Consultado por última vez el 14 de marzo de 2018).

más fuerza la idea de que estos no son susceptibles de apropiación. Sería casi como un enriquecimiento injusto de los científicos el que los pagadores de la obtención de los datos tuvieran que volver a pagar para hacer uso de los mismos.

Existe, además, una corriente más moderna que está cuestionando que los datos puedan «pertenecer» a alguien y la posibilidad de proteger los datos con *copyright* o derecho de propiedad intelectual (Jones, 1986), pues consideran que los datos derivados de la investigación subvencionada deben estar en el dominio público. En efecto, esta corriente defiende que los científicos no pueden ser propietarios de los datos por múltiples razones; una de ellas es que las investigaciones científicas realizadas, en su mayoría, están financiadas con fondos públicos. Este hecho hace que, cuando menos, proceda plantearse si puede defenderse que los científicos que investigan financiados con fondos públicos puedan ostentar algún tipo de propiedad o titularidad sobre los datos. De hecho, esto ocurre ya en el ámbito privado, en la medida en que lo que desarrolle o descubra un trabajador en el curso de su trabajo será propiedad del empleador. Sobre estas formas de protección de los datos se ahondará en el siguiente punto: la protección de los datos.

La problemática de la propiedad de los datos se plantea en dos ámbitos. Por una parte, afecta a los resultados obtenidos, es decir, al estudio en sí mismo. Por otra, afecta también a los datos que han servido de base para realizar el análisis. Los datos que sirven de base a la opinión y que permitan revisar la opinión, o llegar a otra distinta, o superponer investigaciones, están ahora digitalizados. Respecto a esto, se plantea si puede recurrirse también a la protección de los derechos de propiedad industrial, es decir, «proteger», cerrar los datos que sirven de base para los estudios, dado que son una descripción de la realidad.

Por lo tanto, sin perjuicio de los derechos de propiedad intelectual, existe una fuerte reivindicación de que los *datos* relacionados con la investigación científica tienen que ser *públicos*, pues, al fin y al cabo, los científicos se limitan a descubrir lo que ya existe. O, al menos, si no son públicos, que se cuestione la susceptibilidad de su apropiación. De hecho, la esencia del Research Data Law es la no susceptibilidad de apropiación ni *ex lege* ni mediante confidencialidad o derecho de propiedad intelectual de base legal o contractual de datos sometidos al estatuto de acceso y reutilización en abierto de todo dato obtenido mediante financiación pública (Alonso García, 2016).

No obstante, el que los datos sean accesibles o públicos no significa necesariamente que cualquiera pueda hacer uso de los mismos (Costello *et al.*, 2013). La propiedad de los datos permite establecer restricciones en el uso de los mismos, mediante el uso de distintas herramientas que después se analizarán con más detalle. Así, se plantea toda una serie de cuestiones relacionadas con esto, por ejemplo: ¿es posible

que el titular, descubridor de los datos, propietario o no de los mismos, restrinja su uso?

En el sentido fáctico el autor puede limitarlos al no publicarlos. Sin embargo, la publicación no es suficiente para considerar que no quiera establecer restricciones. En efecto, aún en los supuestos en los que los datos no sean susceptibles de protección con sujeción a los regímenes que se analizarán en el punto relacionado con la protección, lo que sí puede decirse es que los datos son propiedad de la persona que los ha ordenado y descubierto, y que este titular de los datos, con las salvedades apuntadas en el caso de investigaciones financiadas con fondos públicos, tiene la potestad de decidir compartirlos o no.

De forma mucho más reciente, se señala la nueva problemática que la era del conocimiento plantea en relación con la propiedad relacionada con los datos (Hoeren, 2014). Parece que este nuevo entorno cuestiona toda la construcción de la propiedad heredada del derecho romano.

Aunque, en un primer momento, la propiedad de los datos parecía estar ligada a la propiedad de los soportes en los que dichos datos pueden leerse (*hardware, flashdrives, etc.*), en el mundo tecnológico actual, esto no tiene ya sentido (Hoeren, 2014). Hoy, los datos ya no están almacenados en sitios concretos y por ello se ha difuminado el concepto de propiedad de los datos. No queda claro si los datos son tangibles o intangibles ni quién puede ser el propietario de los mismos, suponiendo que tenga que existir un propietario de los datos.

De esta manera, el titular de los datos puede decidir imponer restricciones a su uso y reutilización y, por lo tanto, establecer barreras para su utilización y reutilización. Además, se plantea otra reflexión más: si los datos que se han extraído para alcanzar las conclusiones son los correctos o si no lo son, y si, con esos mismos datos, podría haberse llegado a otra conclusión.

De alguna manera, las investigaciones o los investigadores quedan más expuestos si comparten los datos utilizados. Existe una creencia errónea muy extendida en relación con los datos científicos o de biodiversidad que es que los datos publicados son públicos y pueden utilizarse y reutilizarse libremente, sin restricciones. No siempre es así, por lo que será necesario analizar si existen cláusulas que, de hecho, establezcan restricciones en su uso.

Como consecuencia de ello —como se analizará después en el punto relacionado con la interoperabilidad— es clave la decisión que tome el propietario, el titular, el generador de los datos o el nombre que quiera dársele, en relación con los datos, y que no imponga restricciones de ninguna índole en lo que respecta a los datos. Esto se debe a que, incluso en el caso de datos considerados públicos o abiertos, pueden imponerse restricciones al uso, es decir, cesión de los mismos mediante uso de

cláusulas contractuales. Por último, procede abordar la problemática de las formas de protección de los datos.

3.1.3.2. La protección de los datos

El dato, con la tecnología de hoy, es un elemento fundamental; no puede negarse su valor. De hecho, para las compañías es, sin lugar a dudas, un activo susceptible de protección y explotación mediante la aplicación de licencias o de venta a terceros (Glazer *et al.*, 2013).

Los datos son el nuevo oro. Las empresas viven de la explotación de datos obtenidos de distintas fuentes: Facebook, LinkedIn y un sinnúmero de redes y aplicaciones. La aparición del Internet de las cosas, de la domótica y demás están, sin duda, contribuyendo a la aceleración del proceso.

La teoría del *copyright* o derechos de autor y de los derechos de propiedad industrial ha sido muy útil en el pasado para incentivar la innovación (Elliot, 2005). En efecto, históricamente han sido herramientas para los creadores de obras, los innovadores, para explotar de forma exclusiva sus derechos sin competencia. De alguna manera, se premiaba al descubridor o creador permitiéndole recuperar la inversión e impidiendo a terceros copiar su trabajo. Transcurrido un periodo de tiempo, esos derechos de explotación exclusiva desaparecen y cualquiera puede competir teniendo acceso a los datos que, una vez acabados los derechos, quedan en el dominio público.

The Royal Society Intellectual Property Group (2003) analiza, con acierto, la tensión existente entre estas dos corrientes que, a fin de cuentas, puede resumirse en lo que sigue: por un lado, la necesidad de promover la investigación protegiendo de alguna forma los derechos de propiedad industrial y permitiendo su explotación y, por otro, la necesidad de que los derechos de propiedad industrial no constituyan una barrera o que de cualquier forma retrasen el desarrollo de la ciencia al representar barreras. Existe una clara tensión entre el beneficio privado y el bien común (The Royal Society Intellectual Property Group, 2003).

Normalmente, esta tensión se resuelve en el mercado; por lo general, se aplican derechos de propiedad industrial —lo cual puede llevar a una ocultación de datos científicos—. ¿Tiene sentido ahora el exceso de protección? Parece que con la proliferación de las tecnologías, la imposibilidad de utilizar datos y la posibilidad de protegerlos puede tener un alto impacto en el avance de la ciencia.

En efecto, en un principio, el sistema de patentes perseguía lograr que se hicieran públicas las invenciones con la finalidad de construir nuevo conocimiento a partir de ellas. Para ello, otorgaba un derecho exclusivo de explotación (monopolio) sobre un

descubrimiento innovador con la finalidad de reconocer a un inventor su esfuerzo de innovación premiándole con unos ingresos extraordinarios durante un periodo de tiempo.

Sin embargo, este sistema que otrora nos permitiera avanzar, hoy parece no estar sirviendo como lo hacía y empiezan a escucharse voces que lo matizan. Puede ocurrir que la existencia de una patente frene el crecimiento global en determinados momentos y algunos dicen que este sistema es una losa para el desarrollo de la innovación en la actualidad.

Un ejemplo que podría ilustrar esto fueron los descubrimientos, a finales del XIX y principios del XX, de Nikola Tesla que este nunca patentó y que permitieron realizar grandes avances en ingeniería eléctrica, a diferencia de Thomas Edison, que sí patentó sus inventos.

Otro ejemplo de cómo la inexistencia de las patentes puede promover más «innovación» y desarrollo que la existencia de las mismas es el *jazz*. En efecto, gracias a que no se registraron los acordes principales, pudo desarrollarse todo el género. Por eso, se empieza a plantear cada vez con más fuerza la cuestión de si debe primar el monopolio de explotación (es decir, la defensa de la patente) o bien la necesidad global de crecimiento.

La cuestión parecía zanjada hace tiempo. La patente es un monopolio de explotación que no admite excepciones. No obstante, parece que, en Estados Unidos, se está construyendo, jurisprudencialmente, una alternativa intermedia. En efecto, en algunos supuestos muy concretos, los tribunales han reconocido que puede existir la posibilidad de exigir una licencia obligatoria en unos términos razonables de coste.

La justificación será, en todo caso, la necesidad de esta licencia para no frenar el crecimiento de la ciencia. Lo difícil será justificar en cada caso la posibilidad de que se haga. Si bien es cierto que el mero hecho de que esta figura se esté aceptando hace que las partes flexibilicen sus posturas. Las visiones de Estados Unidos y Europa permanecen muy alejadas en este punto, aunque se profundizará sobre este punto más adelante.

Por consiguiente, de unos años a esta parte, se está consolidando una nueva teoría de los derechos de autor cuyo eje central se ha desplazado. Si antes la justificación de la protección de los derechos estaba en proteger la innovación, parece que ahora se está moviendo hacia la apertura de los datos para compartirlos, con la finalidad de que la ciencia avance más rápido.

Según Glazer *et al.* (2103), la protección de los datos y la concesión de licencias sobre ellos pueden depender de la naturaleza de los datos recabados, pero también de las razones por las cuales se recaban, de la forma en que se agregan.

Prosiguen los mismos autores señalando que, para que pueda darse la protección mediante el uso de acuerdos de confidencialidad y restricciones de uso, los datos y bases de datos de una parte pueden reunir los requisitos necesarios para que puedan protegerse en una o más de las categorías de propiedad intelectual o industrial, en concreto:

- Como secreto comercial (no es objeto de esta tesis, por lo que solo se menciona como una de la posibilidad de proteger los datos).
- De conformidad con la normativa que regula el *copyright* (véase Protección de Derechos de Autor).
- Como una categoría de propiedad intangible (se analiza con más detalle en el apartado relacionado con la propiedad de los datos).
- A través de la protección de patentes de base de datos (como se analizará más adelante en un estudio comparado entre los dos regímenes de protección diferenciados en Estados Unidos y Europa).

Al igual que con otros activos protegidos como propiedad intelectual, para asegurar la máxima protección de sus datos, la empresa debe:

- Asegurar una cadena clara y verificable de título, ya sea que los datos compilados se generen internamente o utilizando fuentes de terceros.
- Considerar el alcance y la naturaleza de la protección que puede garantizarse, en el marco de los respectivos regímenes de propiedad intelectual: (a) mediante la creación de procedimientos internos y la adopción de controles de acceso (lógicos) y de *hardware*, de sistemas y de redes (físicos) de acceso de usuarios centrados en el desarrollo y la obtención de datos y el mantenimiento de la confidencialidad; (b) por contrato, por ejemplo, garantizando que la concesión de licencias de sus datos sea coherente con el plan de protección elegido y que sus acuerdos incluyan salvaguardias adecuadas contra el uso no autorizado o la divulgación.

Esta introducción es una buena manera de contextualizar las cuestiones que se plantean con carácter general en el ámbito de los datos y la construcción que se ha hecho de los datos de las empresas. Sin embargo, el objeto de esta tesis son los

datos de investigación y, en particular, los datos de investigación relacionados con la biodiversidad.

Por lo tanto, a los efectos de esta tesis doctoral, la protección de los datos opera en tres ámbitos: la protección de las bases de datos (estructuras que almacenan los datos) por derechos de propiedad intelectual, la protección de los datos (contenido) y, en la Unión Europea, lo que se conoce como el derecho *sui generis*.

No obstante, no puede dejar de mencionarse el cuestionamiento previo que existe de lo que respecta a la posibilidad de proteger los datos científicos. En efecto, la cuestión de base que muchos autores se plantean es si los datos de biodiversidad pueden quedar protegidos con estas normativas.

Los datos primarios de las investigaciones no son susceptibles de protección utilizando normas de propiedad intelectual o *copyright* (Jones, 1986). Además de la cuestión de la fijación de la idea en un medio que requería la normativa americana está la cuestión de la originalidad. ¿Existe originalidad en «recoger» un dato de la naturaleza? Incluso, en el supuesto de que pudieran protegerse los datos, ¿hasta qué punto puede «evitarse» que otro investigador o científico encuentre los mismos datos y llegue a las mismas conclusiones?

Desde el punto de vista técnico de ajuste a la normativa Jones (1986) declara que existen dificultades para encuadrar este tipo de datos y darles la protección que la norma establece. En efecto, parece que no reúnen los requisitos necesarios puesto que el científico se limita a recoger u observar datos de otro creador: la naturaleza. De alguna manera, las observaciones científicas se limitan a recoger la realidad.

Después de un profundo análisis de las distintas cuestiones que no se estima necesario reproducir aquí, el autor concluye que los datos científicos no son susceptibles de protección y concuerda con los tribunales estadounidenses que defienden la necesidad de limitar la protección de este tipo de datos.

Si bien Jones (1986) reconoce que no se puede aplicar a este tipo de datos la protección de las normas de propiedad intelectual o *copyright*, sí reconoce la necesidad de aplicarles algún tipo de protección, como la regulación de la apropiación indebida, o bien la protección contractual.

En el mismo sentido que Glazer *et al.* (2013), Carroll (2015) hace referencia a los distintos sistemas de protección a los que se puede recurrir. La clave no es solo tener en cuenta el régimen jurídico de protección que ampara los datos, sino también definir quién será el titular de estos derechos, sobre todo en lo que respecta a los datos de investigación, puesto que será al titular de los datos a quien habrá que recurrir a la hora de hacer uso de los mismos.

Existen formas de proteger lo intangible y que permiten a los científicos proteger sus descubrimientos sin limitar de forma indebida el conocimiento. Los científicos pueden, de hecho, limitar de forma contractual el uso de los datos, por ejemplo: mediante el establecimiento de licencias.

Esto plantea toda una serie de cuestiones en lo que respecta a la reutilización de los datos, que en sí misma puede llevar aparejadas restricciones al uso de los datos. Otra de las grandes dificultades es que, a nivel mundial, las regulaciones son heterogéneas, siendo así que la ciencia es cada vez más global. Las normas territoriales que limitan la expansión de la ciencia están también, de alguna manera, actuando como una barrera a su expansión.

Después de esta contextualización de cuestiones previas: la definición de dato y las cuestiones relacionadas tanto con la propiedad como con la protección de los datos, se va a analizar estos tres ámbitos haciendo un análisis comparado entre la Unión Europea y Estados Unidos.

3.2. Dos enfoques distintos: Europa frente a Estados Unidos

En efecto, los planteamientos difieren en ambos regímenes. Se va a ir analizando cada una de las protecciones y cómo este tema se resuelve en Europa y en Estados Unidos.

3.2.1. La protección de las bases de datos

A la hora de abordar las cuestiones relacionadas con la protección de la propia estructura de las bases de datos (el continente) es clave entender claramente lo que son, quiénes son los usuarios de las mismas, quién las construye y la finalidad que tienen (Maurer, 2005); en sentido amplio, por lo tanto, las bases de datos pueden actuar como cintas transportadoras entre los que descubren la información y los que la usan.

El presupuesto para analizar la protección que se les dispensa es definir qué es lo que se entiende por base de datos. La definición europea de base de datos señala que es una recopilación de obras independientes, datos u otros materiales organizados de manera sistemática o metódica y accesibles individualmente por medios electrónicos o por cualesquiera otros (Maurer, 2005).

La cuestión de la protección de las bases de datos está contemplada cada vez con más frecuencia en la agenda de los organismos nacionales, internacionales y no gubernamentales. La justificación puede encontrarse en la proliferación de la industria de la información que se ha convertido en uno de los principales mercados en términos de crecimiento en el mundo industrializado. La industria de las bases de

datos se está convirtiendo en una industria clave para todos los agentes (Estados, órganos no gubernamentales, etc.) (Grosheide, 2002).

Más allá de estas reflexiones, Internet ha supuesto un cambio sustancial en lo que respecta al *copyright* o derechos de propiedad intelectual (Lessig, 2004b). En efecto, esta normativa perseguía regular las copias. Aunque leer un libro, no es copiarlo, Internet aporta un enfoque distinto, pues cualquier acción realizada en Internet es una copia y eso afecta sustancialmente a la definición de *copyright*. Por consiguiente, los propios cambios en el soporte acarrear modificaciones en las aplicaciones del *copyright* que plantean toda una serie de problemas. Esto no es todo, pues ahora los permisos de *copyright* están, de alguna forma, imbuidos dentro de los propios instrumentos, es decir, que las propias máquinas aplican de forma automática las restricciones.

3.2.1.1. El modelo europeo

Antes de la directiva, la protección dentro de la Unión Europea estaba dividida en dos grandes grupos:

- Los países que requerían un umbral de originalidad muy bajo para poder protegerlo (Inglaterra, Irlanda, Holanda).
- Los países que exigían un umbral elevado de originalidad para que fuera susceptible de protegerse (Europa continental).

De alguna manera, la directiva buscó un camino intermedio entre el reducido nivel de protección de unos países y el excesivo de otros, una armonización **mediante** la adopción de niveles de protección equivalentes que permitiera la libre circulación dentro del territorio de la Unión Europea (Schneider, 1998).

Por eso, algunos autores interpretaron que la Unión Europea optó por un modelo agresivo (McManis, 2001) con la creación de un derecho específico para la protección de las bases de datos, de la estructura, después de más de ocho años de discusiones sobre la manera en la que mejorar el gran desequilibrio que existe entre el nivel de inversión en el sector de las bases de datos tanto dentro de los Estados miembros como dentro de la Comunidad.

Lo hizo en la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996. Dicha directiva prohíbe copiar, sin los correspondientes permisos, las bases de datos —de hecho, amenaza a los Estados Unidos, para el supuesto de que adopte unas directrices en una línea distinta— (Maurer, 1999).

La finalidad que esta directiva perseguía era conseguir que se incrementara la cantidad de bases de datos en la Unión Europea (Davison, 2007) y, al mismo tiempo, armonizar los regímenes de aplicación a las mismas.

En sus considerandos, al exponer los motivos en los que se basó para la aprobación de esta directiva, la Unión Europea pone de manifiesto que la protección de las bases de datos en Europa no era ni suficiente ni homogénea,²⁶ y que estas diferencias en la protección de las bases de datos estaban teniendo un impacto negativo en el funcionamiento del mercado interior.²⁷ Se estaban produciendo disfunciones que afectaban a la libre circulación.

Por lo tanto, prosigue la directiva en sus considerandos, estas diferencias deben eliminarse²⁸ dado que, dentro de la Unión Europea, se reconocen distintos derechos de autor a las bases de datos,²⁹ que los derechos de autor son una forma apropiada

²⁶ Considerando primero de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(1) Considerando que, en la actualidad, las bases de datos no están suficientemente protegidas en todos los Estados miembros por la legislación vigente y que, cuando existe, tal protección no es uniforme».

²⁷ Considerando segundo de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(2) Considerando que las diferencias de protección jurídica de las bases de datos en las legislaciones de los Estados miembros inciden de forma directa y negativa en el funcionamiento del mercado interior en lo que se refiere a las bases de datos y, en particular, en la libertad de las personas físicas y jurídicas de suministrar bienes y prestar servicios en el sector de las bases de datos de acceso en línea conforme a un fundamento jurídico armonizado en toda la Comunidad; que dichas diferencias pueden agudizarse a medida que los Estados miembros adopten nuevas disposiciones en un sector que está cobrando una dimensión cada vez más internacional».

²⁸ Considerando tercero de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(3) Considerando que deben suprimirse las diferencias que tienen un efecto distorsionador sobre el funcionamiento del mercado interior y que debe prevenirse la aparición de otras nuevas; que no es preciso eliminar las diferencias que en la actualidad no afectan negativamente al funcionamiento del mercado interior o al desarrollo de un mercado de la información en la Comunidad».

²⁹ Considerando cuarto de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(4) Considerando que en los Estados miembros se reconoce una protección de derechos de autor, bajo diferentes formas, respecto a las bases de datos, de acuerdo con su propia legislación o jurisprudencia, y que estos derechos de propiedad intelectual no armonizados pueden tener como efecto impedir la libre circulación de mercancías y servicios en la Comunidad si en las legislaciones de los Estados miembros subsisten diferencias respecto al alcance y las condiciones de protección de los derechos».

de derechos exclusivos³⁰ y que procede reconocer el trabajo, dinero y tiempo invertidos en la creación de la base de datos.³¹

En efecto, antes de la aprobación de la directiva existía una variedad de normativas dentro de la Unión Europea, como puso de manifiesto el Green Paper on Copyright and the Challenge of Technology (COM, 1988).

Una serie de factores movieron a la Comisión Europea a promover la armonización de la norma reguladora de las bases de datos. La rápida expansión de Internet (US Copyright Office, 1997) creó conciencia sobre el crecimiento exponencial de información, la necesidad de procesarla y la importancia de las bases de datos a la hora de desarrollar un mercado de la información dentro de la Unión Europea. Se pone, además, de manifiesto el desequilibrio entre el nivel de inversión en el sector de las bases de datos.

La regulación de la directiva es de aplicación a todas las compilaciones de datos, con independencia de su soporte (electrónicas o en papel). Este trabajo va a centrarse en el análisis de la situación resultante de la directiva, que establece tres niveles de protección:

- La protección de los datos (contenido).
- La protección de la estructura mediante la propiedad intelectual (continente).
- La protección *sui generis*, que afecta al contenido de la base de datos.

A. La protección del contenido (los datos)

En relación con el contenido de la base de datos, el artículo 3 de la Directiva 96/9/CE dispone que «la protección del derecho de autor de la presente directiva [...] se entenderá *sin perjuicio de los derechos que pudieran subsistir sobre dicho contenido*».

³⁰ Considerando quinto de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(5) Considerando que los derechos de autor constituyen una forma apropiada de derechos exclusivos de los autores de las bases de datos».

³¹ Considerandos séptimo, octavo y noveno de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos: «(7) Considerando que la fabricación de una base de datos requiere una gran inversión en términos de recursos humanos, técnicos y económicos, y que las bases de datos se pueden copiar o se puede acceder a ellas a un coste muy inferior al necesario para crearlas de forma independiente»; «(8) Considerando que la extracción o reutilización no autorizadas del contenido de una base de datos son actos que pueden tener consecuencias graves desde el punto de vista económico y técnico»; «(9) Considerando que las bases de datos constituyen un instrumento de gran valor para el desarrollo del mercado comunitario de la información; que este instrumento es de gran utilidad para otras muchas actividades».

Reconoce, por consiguiente, que el contenido de las bases de datos puede ser susceptible de protegerse.

En realidad, en la protección de las bases de datos lo que realmente interesa es la protección del contenido (Grosheide, 2002). De hecho, prosigue, los datos no adquieren derecho a protección por haber sido incluidos en la base de datos; solo si antes eran susceptibles de protegerse lo serán una vez incluidos en la base de datos.

Son dos derechos perfectamente diferenciados e independientes el uno del otro, la protección del continente (base de datos) y la protección del contenido (datos). Tal vez proceda una reflexión previa para poner de manifiesto que el hecho de que los datos se publiquen o aparezcan en una base de datos no pueden usarse libremente. En efecto, estos datos pueden quedar protegidos con una variedad de regímenes jurídicos.

Una de las posibilidades, sin lugar a dudas, serán los derechos de propiedad intelectual: patentes, derechos sobre dibujos y modelos y derechos de autor, pero, como contenido de una base de datos, parece complicado aplicar el derecho de marcas.

En efecto, la Convención de Berna prevé que los contenidos de las bases de datos, siempre que demuestren originalidad y creatividad, serán susceptibles de ser protegidos en Europa (Harison, 2010). Los derechos de propiedad intelectual, a diferencia de las patentes, no requieren la formalidad de ser registrados.

Los contenidos no están protegidos en todos los casos, sino solo en aquellos en que su creación haya llevado aparejada un esfuerzo intelectual, por lo que no será de aplicación en el caso de que se incluyan hechos, por ejemplo. Adicionalmente, la directiva ha incluido una protección concreta llamada *sui generis* —que se analizará con más detalle después— que afecta a los contenidos de las bases de datos que no sean susceptibles de protección con arreglo a ninguno de los regímenes contemplados en este apartado.

Otra de las formas de protección, que siempre se ha considerado alternativo o complementario al derecho de propiedad intelectual (Erdozain López, 2016) es el que establece la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas (Parlamento Europeo y Consejo de la Unión Europea, 2016).

De conformidad con el considerando primero de esta directiva:

[L]as empresas, así como los organismos de investigación de carácter no comercial, invierten en la obtención, desarrollo y aplicación de conocimientos técnicos (*know how*) e información, que son la moneda de cambio de la economía del conocimiento y proporcionan una ventaja competitiva. Esta

inversión en la generación y aplicación de capital intelectual es un factor determinante para su competitividad y su rendimiento asociado a la innovación en el mercado y, por tanto, para la rentabilidad de sus inversiones, que constituye la motivación subyacente a la investigación y el desarrollo en las empresas. Las empresas utilizan diferentes medios para hacer suyos los resultados de sus actividades asociadas a la innovación cuando optar por la apertura no permite la plena explotación de sus inversiones en investigación e innovación [...]. Otro medio para apropiarse de los resultados de la innovación consiste en proteger el acceso a los conocimientos que son valiosos para la entidad y que no son ampliamente conocidos, y explotarlos. Esos conocimientos técnicos y esa información empresarial de gran valor, que no se han divulgado y que se quieren mantener confidenciales, se conocen con el nombre de *secretos comerciales*.

A los efectos de proteger estos secretos comerciales, las empresas utilizan los mecanismos de la protección de la confidencialidad de los datos. El considerando cuarto prosigue señalando que:

Las empresas innovadoras están cada vez más expuestas a prácticas desleales que persiguen la apropiación indebida de secretos comerciales, como el robo, la copia no autorizada, el espionaje económico o el incumplimiento de los requisitos de confidencialidad, ya sea dentro o fuera del territorio de la Unión.

A pesar de que la protección de los secretos comerciales quedó recogida en el acuerdo sobre los Acuerdos sobre los Aspectos de los Derechos de Propiedad Intelectual (conocidos por sus siglas ADPIC o, en inglés, TRIPS) (*Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio*, 1994), existen divergencias en los regímenes nacionales dentro de la Unión Europea que, por consiguiente, generan también unas diferencias en los niveles de protección jurídica de los secretos comerciales dentro del territorio de la Unión Europea.

La directiva tiene como objetivo armonizar las medidas contra la obtención no autorizada y el posterior uso de secretos comerciales con el fin de mejorar el funcionamiento del mercado interior. Para ello, armoniza el concepto de «secreto comercial» estableciendo como características que debe contener el mismo (Mediano, 2016):

1. Que sea una información secreta en el sentido de que no sea, como cuerpo o en la configuración y reunión precisas de sus componentes, generalmente conocida ni fácilmente accesible para las personas pertenecientes a los círculos en que normalmente se utilice el tipo de información en cuestión.
2. Que tenga un valor comercial por su carácter secreto.

3. Que haya sido objeto de medidas razonables en las circunstancias, para mantenerla secreta, tomadas por la persona que legalmente ejerza su control.

Existe otro régimen de protección que también puede ser de aplicación a los contenidos de las bases de datos: la normativa de protección de datos de carácter personal.

En efecto, existe una normativa que regula todos estos aspectos (Parlamento Europeo y Consejo de la Unión Europea, 2016) donde se regula de forma detallada el uso de los datos de carácter personal gestionados tanto de forma manual, como automática y que, por ende, también será de aplicación a las bases de datos. Se da una definición de lo que son los datos de carácter personal y de cómo se tienen que gestionar este tipo de datos. No podía dejarse de mencionar este régimen de protección de contenidos, aunque solo brevemente, puesto que los datos de carácter personal no son objeto de este trabajo y profundizar sobras las cuestiones relacionadas con ello podría dar lugar a otra tesis completa sobre el tema.

El dato, como ya se ha puesto de manifiesto con anterioridad, es el nuevo oro de la presente era digital. Por lo tanto, parece complicado defender que, cuando menos en algunos supuestos, los datos no sean susceptibles de protección. Dado que el objeto de este trabajo son los datos científicos, el análisis sobre la posibilidad de proteger los datos se centrará en este tipo.

Aunque los científicos pueden tener interés en proteger los datos de investigación de los competidores, sobre todo antes de que sean publicados, o en proteger los datos para acreditar que han sido los primeros en descubrirlos y poder utilizarlos para pedir subvenciones (Jones, 1986), esto es cada vez más cuestionable.

Esta práctica hace necesario buscar nuevas fórmulas en lo que respecta a los datos de investigación y científicos. Sin duda, el origen público o privado de los datos es importante a la hora de evaluar y establecer los sistemas de protección que les puedan ser de aplicación. No será lo mismo que los datos se hayan obtenido en investigaciones realizadas en el marco de distintos tipos de subvenciones con fondos públicos que el hecho de que los hayan obtenido empresas privadas en el desarrollo de su negocio.

La relación entre los derechos de propiedad intelectual y la difusión del conocimiento científico ha sido durante mucho tiempo una preocupación (Murray y Stern, 2006). Una cuestión clave de la discusión es si puede protegerse la investigación básica no comercial.

B. La protección de la estructura (el continente)

La estructura de la base de datos quedó protegida por razón de la inversión que suponía (Parlamento Europeo y Consejo de la Unión Europea, 1996). De hecho, con arreglo a los ADPIC antes mencionados (*Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio*, 1994) se planteaba si, por los asuntos de que se trata o por la forma en que están dispuestos, constituían una creación intelectual propia de su autor.

De conformidad con el artículo 3.1 de la directiva:

Las bases de datos que por la selección o la disposición de su contenido constituyan una *creación intelectual de su autor* estarán protegidas, como tal creación, por los derechos de autor. No serán de aplicación otros criterios para determinar si tales bases de datos son susceptibles de dicha protección [el resaltado es nuestro].

Según recoge Grosheide (2002), el fabricante de una base de datos, de la estructura en sí, dispone de un conjunto de derechos exclusivos (actos sujetos a restricciones como la reproducción, transformación, distribución, etc.).

Esta protección establecía, en primer lugar, un estándar de originalidad. Además, recogía una serie de «actos sujetos a restricciones» y de excepciones a estos. El usuario legítimo de una base de datos debía poder efectuar todas las actividades necesarias para poder trabajar con la propia base de datos, dentro de ciertos límites.

La directiva estableció que los actos sujetos a restricciones realizados con la finalidad de acceder a los contenidos de la base de datos no quedasen sujetos a autorización. De igual manera, la directiva preveía la posibilidad de que los Estados miembros establecieran limitaciones sobre los actos sujetos a restricciones como (US Copyright Office, 1997):

- La reproducción para fines privados.
- El uso solo a efectos de enseñar o de realizar investigaciones científicas.
- El uso con fines de seguridad pública.
- Otras excepciones que puedan existir en las leyes nacionales.

En cualquier caso, detrás de estas excepciones, subyacía el principio de que no pueden causarse perjuicios económicos. De alguna manera, la directiva recoge el principio de protección conocido en Estados Unidos como *sweat of the brow*, cuyo análisis se realizará en el punto del presente trabajo relativo a Estados Unidos (Hugenholtz, 2001). Por lo tanto, es susceptible de protección por la mera inversión realizada, aunque no se den referencias de la misma.

C. Los derechos *sui generis*

Estos derechos se recogen en la Directiva (CE) n.º 96/9 del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos (cap. 3). Cuando la Comunidad Europea emitió su primera propuesta sobre la protección jurídica de las bases de datos en 1992, la memoria explicativa adjunta hacía expresa referencia a la Decisión Feist (Feist Publications Inc. v. Rural Tel. Serv. Co, 1991). Señalaba que las bases de datos electrónicas que no cumplieran con la prueba de originalidad quedarían excluidas de la protección, independientemente de la habilidad, mano de obra, esfuerzo o inversión financiera que hubiera en la creación de las bases de datos electrónicas (Grosheide, 2002).

Sin embargo, la redacción de 1996 introdujo el derecho *sui generis*. Como ya se ha apuntado, este derecho recogido en la directiva europea de bases de datos es el resultado del compromiso entre los países en los que existe el derecho de autor, liderados por Francia y adheridos a una norma de «creación intelectual» estándar y los países de derecho anglosajón como Inglaterra e Irlanda que también atribuyen protección de *copyright* a las compilaciones que requieren una inversión sustancial al amparo de la doctrina del *sweat of the brow* (Cardinale, 2007).

Una de las razones para que esta directiva se aprobara fue la presión de Francia para protegerse de las posibles prácticas depredadoras de los productores de Estados Unidos.

De hecho, en el año 1996 la Unión Europea (con la US Trademark Office) presentó una propuesta a la Organización Mundial de la Propiedad Intelectual (en adelante, OMPI), para firmar un tratado en relación con la protección jurídica de las bases de datos. La idea era introducir en el tratado de la OMPI el concepto de la protección *sui generis* (Band y Kono, 2001).

De alguna manera, se ha venido considerando que la protección de obras creativas e invenciones es una forma de promover su creación, aunque existen grandes defensores de que estos incentivos económicos no eran necesarios; fue un tema muy controvertido.

En el seno de la OMPI, se realizaron una multitud de estudios, en relación con el impacto de la protección de las bases de datos «no originales», que son aquellas que no satisfacen el criterio de originalidad de los principios de derecho de autor reconocidos en el Convenio de Berna (Anon. 1886) o en el tratado de la OMPI sobre los derechos de autor (Braunstein, 2002).

La finalidad de la protección de la estructura de las bases de datos, a fin de cuentas, no era otra que preservar el coste de la inversión realizado, protegiéndola de la copia

que tiene un coste muy bajo. En el otro lado, estaba la necesidad de hacer los datos accesibles a la generalidad.

Con estos presupuestos, se analizará el desarrollo en la normativa europea. En efecto, la directiva, además de la armonización del derecho de propiedad intelectual europeo en relación con la compilación de las bases de datos, obligaba también a sus miembros a crear una nueva forma de protección de propiedad intelectual en relación con el acceso a los contenidos de las bases de datos, los llamados derechos *sui generis* (McManis, 2001). El artículo 7 los definió poniendo de manifiesto que:

Los Estados miembros dispondrán que el fabricante de la base de datos pueda prohibir la extracción o reutilización de la totalidad o de una parte sustancial del contenido de esta, evaluada cualitativa o cuantitativamente, cuando la obtención, la verificación o la presentación de dicho contenido representen una inversión sustancial desde el punto de vista cuantitativo o cualitativo.

La Unión Europea parte del concepto de que crear una base de datos representa un coste elevado, mientras que copiar o reproducir la base de datos no lo tiene (Schneider, 1998). Habla de «inversión sustancial», término que no queda claramente definido. Muchos autores se preguntan por el contenido de este término (Derclaye, 2005). De alguna manera, lo que se pretendía era proteger la inversión, aunque al mismo tiempo el riesgo era consustancial al término inversión, por lo que puede parecer contradictorio (Schneider, 1998).

No cabe duda de que la denominación *sui generis* es muy acertada, habida cuenta de que es un derecho único y que no hay precedentes en el derecho internacional (Cardinale, 2007). En ninguna otra jurisdicción se hacen distinciones entre las bases de datos creativas y las no creativas. Además, la directiva recogía una cláusula de reciprocidad en la aplicación de la normativa, muy controvertida y cuya aplicabilidad se ha cuestionado en una multiplicidad de ocasiones.

De acuerdo con este régimen *sui generis*, el fabricante de una base de datos, persona física o jurídica, podía prohibir la extracción o la reutilización no autorizada del contenido de una base de datos (McManis, 2001).

La directiva justificó la introducción de esta protección en su considerando 6 estableciendo que «se precisan unas medidas que impidan la extracción o reutilización no autorizadas del contenido de una base de datos a falta de un régimen armonizado relativo a la competencia desleal o de la correspondiente jurisprudencia». Si bien, en los previos de la directiva, los Estados miembros quisieron poner énfasis en la competencia desleal; la falta de armonización de los regímenes aplicables en los distintos Estados hizo que cada vez más desplazaran este interés a una protección entroncada en la propiedad intelectual.

Las principales características del derecho *sui generis* que recoge la Directiva son las siguientes (US Copyright Office, 1997):

1. Protección de la *inversión sustancial*. La titularidad de los derechos *sui generis* la tiene el fabricante «de la totalidad o de una parte sustancial del contenido de esta, evaluada cualitativa o cuantitativamente, cuando la obtención, la verificación o la presentación de dicho contenido representen una inversión sustancial desde el punto de vista cuantitativo o cualitativo».
2. Protección contra los *actos de extracción* (Dir. 96/9 del Parlamento y del Consejo, de 8 de diciembre de 2015, art. 2.a) y *reutilización* (Dir. 96/9 del Parlamento y del Consejo, de 8 de diciembre de 2015, art. 2.b).
3. Las *partes no sustanciales* quedan *excluidas* de la protección (Dir. 96/9 del Parlamento y del Consejo, de 8 de diciembre de 2015, art. 8.1).
4. *Excepción de ciertos usos* (Dir. 96/9 del Parlamento y del Consejo, de 8 de diciembre de 2015, art. 9) como puede ser el uso privado, el uso con fines de enseñanza o de investigación científica. Un usuario legítimo puede extraer o reutilizar sin autorización partes no sustanciales del contenido de la base. No obstante, no puede efectuar actividades que perjudiquen de forma injustificada los intereses legítimos del fabricante de la base o del proveedor de las obras o prestaciones contenidas en la misma.

Los derechos *sui generis*, en tanto que derechos económicos, pueden transferirse, cederse o darse en licencia contractual.

La protección frente a la extracción no autorizada del contenido de una base de datos se concede por un periodo de quince años a partir de la fecha de finalización de la creación. Davison (2007) señala que en algunos aspectos —como puede ser la duración de la protección al amparo del derecho de bases de datos, que está teóricamente limitado a quince años, pero en la práctica es prácticamente perpetua— el derecho de protección va más allá del *sweat of the brow*. Siempre que se actualice la base de datos y la actualización pueda suponer una inversión sustancial en lo que respecta a la información incluida en la misma, el periodo de protección podrá renovarse. El nuevo periodo de protección afecta no solo a la actualización, sino también a todos los datos parte del proyecto inicial, incluso en el supuesto de que estos no se hayan actualizado. La finalidad era conseguir un mayor número de bases de datos en la Unión Europea.

A primera vista, puede considerarse un derecho muy parecido al *copyright* del *sweat of the brow* que se estudiará con detalle en el apartado dedicado al modelo americano. Sin embargo, en algunas cosas, va incluso un paso más allá (Davison, 2007).

Además de la protección «ilimitada» ya analizada, Davison (2007) habla también de la vulneración del derecho que representa la apropiación de una parte sustancial de la base de datos, cuantitativa o cualitativamente. La introducción de las cuestiones cualitativas en la protección el *sweat of the brow* plantea algunas cuestiones alarmantes como podía ser el hecho de que los propietarios de las bases de datos buscasen la protección de uno o más elementos de las mismas con base en que estas sean cualitativamente relevantes. Por lo tanto, se puede buscar la protección de temas cuantitativamente poco significativos usando la relevancia cualitativa.

Existen muy pocas excepciones, incluso menos que para el supuesto de *copyright* recogido en el derecho estadounidense que se analizará después. La principal excepción es para la extracción de imágenes que se usen con fines educativos y científicos, siempre que se indique la fuente y tengan fines no comerciales. Es una excepción construida con conceptos jurídicos indeterminados y, por lo tanto, resulta complicado interpretar lo que significa investigación científica o lo que son usos no comerciales. Algunos de los países de la Unión Europea (Italia, Francia o Irlanda) no han recogido estas excepciones y no tienen, por ejemplo, derecho de *fair use*.

Por estas y por otras cuestiones la protección de las bases de datos es superior en la Unión Europea que en Estados Unidos. Este distinto nivel de protección ha llevado a algunos autores a plantearse si es aconsejable establecer las bases de datos en Europa (Bovemberg, 2000).

La protección contra la extracción o la reutilización no autorizada se concede a las bases de datos cuyo fabricante sea ciudadano residente en la Unión Europea o una sociedad o empresa cuya sede oficial, administración central o establecimiento principal se encuentren en territorio comunitario.

Algunos autores como Lubens (2003) se plantean si el régimen de protección europeo es el adecuado poniendo de manifiesto que es posible que partidarios de los usos libres de Internet se vean amenazados como consecuencia de esta regulación. Algunos usuarios de Internet auguran que las bases de datos terminarán almacenándose fuera de la Unión Europea para evitar la aplicación de estas restricciones. Tal vez proceda mencionar esto, habida cuenta de la nueva regulación del *text and data mining* que se está planteando aprobar la Unión Europea y que, como después se verá, no hace más que reforzar esta idea.

La naturaleza del derecho de bases de datos y los derechos exclusivos del propietario del *copyright* tienen un contenido muy similar. La directiva exigía garantizar un derecho de extracción y un derecho de reutilización. La extracción puede definirse como la transferencia permanente o temporal de todo o parte del contenido a otro soporte por cualquier medio o de cualquier forma. Se trata en realidad del derecho de reproducción. La reutilización se define como cualquier

modo de poner a disposición del público todo o una parte sustancial de los contenidos de la base de datos mediante la distribución de copias, *renting*, *on-line* o mediante cualquier otra forma de transmisión (Davison, 2007).

Por lo tanto, el derecho *sui generis* protege a cualquiera que haga una inversión sustancial de extracciones no autorizadas, y de utilizaciones y reutilizaciones de toda o de una parte sustancial de la base de datos (Reichman y Samuelson, 1997).

3.2.1.2. El modelo estadounidense (US Copyright Office, 1997)

A diferencia de lo que ocurre en Europa, en Estados Unidos no existe una norma específica para proteger las bases de datos. Debe recurrirse a la regulación del *copyright* siempre y cuando reúna las condiciones para considerarse una compilación.

La National Commission on New Technological Uses of Copyrighted Works (CONTU por sus siglas en inglés), en un informe en 1978, declaró que las bases de datos computarizadas podían quedar incluidas dentro de la protección del *copyright* como compilaciones (Duke, 2001). Procede analizar, cuál es la justificación para que una base de datos pueda quedar amparada bajo la protección del *copyright* y qué partes pueden quedar protegidas y cuáles no.

Históricamente, las políticas estadounidenses de información básica establecen que los hechos son de dominio público, por lo tanto, no son susceptibles de apropiación (Band y Kono, 2001). Esto permite extraer datos de una compilación existente para reutilizarlos en otra compilación. Los hechos son los pilares del conocimiento, por lo que cualquiera puede utilizarlos y reutilizarlos (Band y Kono, 2001).

Sin embargo, el que elabora una base de datos puede proteger su inversión en la recogida de datos de cuatro formas diferentes. En primer lugar, puede confiar en el *copyright* (sobre este punto se profundizará un poco más adelante). El *copyright* protege la selección original, la coordinación y la estructura de los hechos en una compilación, pero no los propios hechos. El *copyright* protege la expresión de las ideas y no la idea en sí misma (Broussard, 2007). Por eso, por lo general, el *copyright* protege de la copia completa de una base de datos, que tiene como mínimo un cierto nivel de expresión original, pero no la extracción y reutilización de los hechos individuales (Band y Kono, 2001).

En segundo lugar, puede ampararse en el derecho contractual. Muchas de las bases de datos, en particular las que están en línea, se distribuyen al amparo de acuerdos de licencia de conformidad con los cuales, el licenciataria, usuario, está de acuerdo con redistribuir la información (Band y Kono, 2001).

En tercer lugar, puede confiar en la ley estatal común de apropiación indebida. De conformidad con esta doctrina, el que recoge los datos puede impedir a los

competidores que copien cualquier información sensible. Además, cada vez con más frecuencia, los tribunales reconocen «accesos indebidos» cuando se extrae información de sitios web sin autorización (Band y Kono, 2001).

En cuarto lugar, puede confiar en medidas técnicas, que son particularmente efectivas en lo que respecta a las bases de datos en línea, en los casos en los que se pueda limitar el acceso de los usuarios a relativamente pequeñas cantidades de información de una vez. Estas limitaciones impiden que se copie la totalidad de la base de datos. Además, las medidas tecnológicas ahora reciben protección legal al amparo de la Digital Millennium Copyright Act (DMCA) (Anon., 1978) (Band y Kono, 2001).

Durante el siglo XIX, dos ideas fundamentales sirvieron de justificación para la protección de las bases de datos. La primera fue la conciencia del esfuerzo y la inversión realizados por la persona que recopila los datos, que se conoce como la doctrina *sweat of the brow* (sudor de la frente) (*Sweat of the brow*, s. f).³² La segunda de las ideas sobre las que se fundamenta la protección de las bases de datos fue el reconocimiento del criterio del autor en la elaboración y la selección de la información a incluir en la base de datos.

No obstante, este reconocimiento del «sudor de la frente» y de la selección de la información a incluir en la base de datos deja de ser la clave en lo que respecta a la protección de las bases de datos. Se produce un «tímido cambio» en la legislación estadounidense en la ley de *copyright* de 1976³³ en la que aparece la necesidad de que exista originalidad. Esta regulación va sufriendo diversas modificaciones. De conformidad con la ley de *copyright* estadounidense, una base de datos es una compilación, un trabajo constituido por una colección y recopilación de materiales o datos ya existentes.³⁴

Las compilaciones son una de las formas de autoría protegidas en Estados Unidos desde hace más tiempo; empezaron a protegerse en el siglo XVIII y comenzaron con la protección de los libros.

³² De conformidad con esta doctrina un autor obtiene derechos por el mero hecho de elaborar un trabajo, una base de datos o directorio. No se exige ni creatividad ni originalidad. En efecto, se le reconoce este derecho de autor por el esfuerzo y por la inversión realizada. Por lo tanto, si cualquier persona quiere hacer uso de ese trabajo debe volver a realizarlo de cero y no puede utilizar el trabajo anterior en la elaboración de su directorio.

³³ USC art. 101: «A ‘compilation’ is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship. The term ‘compilation’ includes collective works». Disponible en inglés en <http://www.copyright.gov/title17/92chap1.html#101>.

³⁴ Traducción no oficial.

En 1991, en la decisión tomada en el caso en *Feist Publications v. Rural Telephone Service Co.* (1991) el Tribunal Supremo estadounidense revisa la doctrina «del sudor de la frente» y considera que poder proteger una base de datos con los derechos de autor exige un cierto grado de creatividad, aunque sea pequeño. Por eso, en el caso que les ocupa lo que se discute es la aplicación de la protección derivada del *copyright* a un directorio telefónico; el Tribunal considera que esta forma de organizar la información carece de cualquier originalidad y que, por tanto, no puede protegerse con arreglo al *copyright*. De hecho, el Tribunal Supremo dice que se trata de una forma obvia de organizar la información y que Rural se limita a listar por orden alfabético, a todos sus clientes.

Por lo tanto, la decisión *Feist* (*Feist Publications Inc. v. Rural Tel. Serv. Co.*, 1991) marca un cambio de criterio en Estados Unidos en lo que respecta a la posibilidad de proteger las bases de datos con *copyright*. Si antes de esta decisión cualquier estructura era susceptible de protección aplicando la norma del *sweat of the brow*, ahora se da un paso más y se exige una «cierta» originalidad.

En efecto, una pequeña dosis de originalidad será suficiente para que la base de datos sea susceptible de protección al amparo de los derechos de autor. Por lo tanto, la clave desde *Feist* respecto a la posibilidad de proteger las compilaciones no va a encontrar su fundamento en el esfuerzo ni en la inversión (doctrina del *sweat of the brow*), sino en una cierta creatividad u originalidad en su disposición.

Los datos primarios que sirven en la compilación no serán protegibles en sí mismos, aunque algunos autores como Resnik (2003) se plantean si es necesario incrementar la protección dada, acercándola al planteamiento europeo.

Un análisis sobre cómo funciona hoy el mercado de las bases de datos será una ayuda para comprender mejor la evolución regulatoria habida. Las formas que se usan ahora para proteger las bases de datos y las licencias de uso han cambiado. En efecto, hoy, los propietarios de bases de datos buscan nuevos mecanismos para proteger sus creaciones y son tres las formas de hacerlo que han proliferado.

La primera es incluir datos susceptibles de obtener protección de *copyright*.³⁵ Si bien es cierto que la inclusión de datos sujetos a protección de derechos de autor dificulta la copia directa de la base de datos, sigue siendo posible extraer los datos de la misma que no estén protegidos por el *copyright* y, con ellos, construir una base de datos distinta.

³⁵ *Copyright* y derecho de autor no recogen los mismos conceptos. *Copyright* es una figura de derecho anglosajón, que incluye el derecho de copia y, por lo tanto, se refiere a los derechos patrimoniales de explotación del autor. Mientras que el concepto de derecho de autor, de derecho continental, hace referencia a los derechos morales que no se reconocen en el derecho anglosajón.

La norma estadounidense (sección 102 b) exige un cierto grado de originalidad para que puedan protegerse los distintos elementos y, además, señala que no estarán amparados en el *copyright* las ideas, procedimientos, procesos, sistemas, métodos de operación conceptos, principios o descubrimientos, con independencia de la forma en la que se describa, explique, ilustre o que se le dé (United States, 1976).

En la sección 103 añade que tampoco quedan amparados en la protección del *copyright* los hechos, aunque su selección y arreglo pueden quedar protegidos (Duke, 2001). Esto es consistente con la idea de que los hechos tienen que permanecer en el dominio público.

La segunda manera es incrementar la creatividad en lo que respecta a la estructura de la base de datos y a la forma de organizar los datos. Sin embargo, es importante hacer una buena estructura porque el realizar una más creativa puede crear barreras al acceso a la información haciendo más difícil el manejo de la base de datos. Además, la mayor creatividad exige, evidentemente, una mayor inversión y un mayor coste a la hora de hacer la base de datos.

La tercera es el recurso a las condiciones contractuales para evitar que se copie la base de datos (US Copyright Office, 1997). Dentro de ello existen distintas formas. Por una parte, mediante la inclusión de unas condiciones de uso. Otra posible alternativa es recurrir a una suscripción con un coste periódico (mensual, trimestral, etc.) para regular el acceso a los contenidos, o bien contratando tiempos de uso. Existen excepciones de uso, dando posibilidad de suscripciones gratuitas en el supuesto de que el acceso sea para investigación y no para usos comerciales.

Dentro de este apartado, quedan también incluidas las barreras tecnológicas que los propietarios de las bases de datos incluyen con la finalidad de hacer más difícil el acceso a la base de datos, es decir, la extracción de los datos de la misma para incluirlos en otra base de datos o la copia de la propia base de datos.

Existe todo un debate sobre si es más adecuada la regulación vigente en Europa o la que se dicta en Estados Unidos. De hecho, en el 106th Congreso de Estados Unidos, se plantearon una serie de iniciativas en relación con la protección de las bases de datos.

La primera fue la Collections of Information Antipiracy Act (Anon., 1999), que persigue que una persona no pueda poner a disposición de otros una parte sustancial de una colección de información recopilada o mantenida por otra persona. La segunda, la Consumer and Investor Access to Information Act (Anon., 1999) persigue fundamentalmente a las bases de datos parasitarias de otras, en concreto, prohíbe entregar a nadie un duplicado de la base de datos para competir con la original (Band y Kono, 2001).

Otra de las cuestiones que se plantearon en este Congreso, celebrado en 1999-2001, es si la protección de las bases de datos en Estados Unidos debe asimilarse a la que existe en la Unión Europea desde 1996, con la aprobación de la directiva.

Sin embargo, el equilibrio es clave. La información es el oxígeno de la nueva economía, y cambios drásticos en las políticas de información pueden obstaculizar la expansión de Internet, dificultar la investigación y ahogar la innovación y el emprendimiento (Band y Kono, 2001).

En efecto, la discusión que ahora se plantea es si estos regímenes son los adecuados para hacer frente a la nueva revolución digital en la que los datos son cada vez más accesibles desde el punto de vista técnico, aunque no necesariamente desde el punto de vista jurídico. Por lo tanto, cabe preguntarse si estas regulaciones están contribuyendo a la multiplicación del conocimiento o, por el contrario, frenándolo.

En efecto, el *copyright* engloba una serie de derechos exclusivos atribuidos a los autores para controlar la mayor parte de las reutilizaciones que de sus trabajos puedan hacerse, sin su permiso y con ciertas limitaciones y excepciones. De hecho, lo que justifica la teoría del *copyright* es que los autores usarán esos derechos que ostentan, bien para proceder a publicar lo que han escrito directamente, bien para que lo publique un tercero remunerando al autor. Los autores pueden reservarse el *copyright* y dar una licencia no exclusiva al editor que lo publique (Carroll, 2011).

Sin embargo, con Internet muchas cosas se modifican. Si bien es extremadamente importante que la regulación del *copyright* continúe promoviendo el progreso, tanto en el arte como en las ciencias, y remunerando la creatividad de los autores, al mismo tiempo, hay que reconocer que estas regulaciones pueden ahogar el progreso con una sobreprotección (Blanke, 2002).

Los modelos de acceso abierto, como se verá después, usan este modelo de permisos con la finalidad de garantizar a los lectores una amplia reutilización de derechos para promover una amplia difusión, republicación y reutilización de los artículos. Uno de los medios para realizar esto es el uso de licencias y, en este sentido, nacen algunas fórmulas prerredactadas que después se analizarán. No obstante, eso no quita que el uso de las licencias sea necesario en determinados casos, estableciendo restricciones al uso de los datos. En efecto, una licencia es un contrato vinculante en el cual el proveedor de la base de datos y el usuario acuerdan de conformidad con los principios del derecho contractual (Onsrud y López, 1998). En este contexto, nace todo el concepto de acceso abierto, ¿es el primer paso para la interoperabilidad global?

A. Los datos generados por agencias gubernamentales

En Estados Unidos la Copyright Act de 1976 (Estados Unidos, 1976) prohíbe al Gobierno federal proteger sus publicaciones.³⁶ A resultas de ella, los datos y la información que hayan tenido origen en un proyecto del gobierno entrarán en el dominio público transcurrido un año, sin que a este pueda aplicarse ninguna clase de derecho de propiedad.

Así, la National Science Foundation tiene un programa de subvenciones que también supedita a la publicación de los datos (National Science Foundation, 1995; National Science Foundation, Office of Cyberinfrastructure, Directorate for Computer & Information Science & Engineering, 2008). Además, desde enero de 2011, (Kvalheim y Kvamme, 2014) la National Science Foundation empezó a exigir un plan de gestión de datos de todas las nuevas propuestas para asegurarse de que los investigadores tienen planificado mantener los datos durante la investigación, aunque también después, con la finalidad de maximizar el intercambio de datos. Las propuestas que se presenten o deban presentarse a partir del 18 de enero de 2011 deberán incluir un documento adicional de no más de dos páginas, que recoja el llamado Plan de Gestión de Datos (National Science Foundation, 2010).

Este documento debe describir la propuesta para cumplir con la política de la National Science Foundation en cuestiones de diseminación e intercambio de los resultados de la investigación (National Science Foundation, 2013), así como enunciar las fuentes y recursos físicos y virtuales utilizables para la efectiva preservación y el almacenamiento de los datos resultantes de la investigación. Este plan puede incluir instalaciones de terceros y repositorios.

Es decisión del investigador especificar y describir los recursos cibernéticos que deban usarse para la preservación y almacenamiento de datos. Solo hay una especificación adicional que hacer en lo que respecta a la diseminación e intercambio de datos: todos los proyectos de investigación subvencionados deben, de forma obligatoria, compartir los resultados, los datos, las muestras, las colecciones físicas y cualesquiera otros materiales de apoyo obtenidos o reunidos en el curso del desarrollo del trabajo, con un coste no superior al incremental y en un plazo de tiempo razonable. Además, se espera que los subvencionados promuevan y faciliten ese intercambio de datos.

³⁶ Copyright Act de 1976 de Estados Unidos, artículo 105: «Subject matter of copyright: United States Government works. Copyright protection under this title is not available for any work of the United States Government, but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise». Se puede consultar en www.copyright.gov/title17/92chap1.html#105. (Consultado por última vez el 25 de mayo de 2017).

El Directorado de Evaluación de Ciencias Sociales, Económicas y de Comportamiento de la National Science Foundation formula una política de archivo de datos todavía más específica que contiene directrices en función de las categorías de datos. Es un ejemplo de políticas con requerimientos flexibles (National Science Foundation, s. f.).

En esa misma línea, el National Institute of Health estadounidense (NIH) tiene requisitos concretos de estructuras de preservación de datos en lo que respecta a las publicaciones científicas, de manera que se exige que todos los investigadores subvencionados presenten una versión final, revisada por pares, del trabajo para publicar en formato electrónico en la National Library of Medicine's PubMed Central.

El National Institute of Health dispone, de hecho, de una política de compartir datos que condiciona la concesión de la subvención a la presentación de un plan de difusión, de compartir los resultados de los estudios que se realicen con la finalidad de que los resultados de un tema puedan servir para adelantar desarrollos de otros (National Institute of Health, 2003).

Como adelantábamos más arriba, conviene aclarar qué es *public domain* o dominio público. Este concepto se define como aquella información que procede de una fuente o que está integrada por un tipo de datos cuyos usos no están restringidos por los derechos de propiedad intelectual (IP) o por cualesquiera otros regímenes legales (Esanu y Uhler, 2003). Se trata de datos que puede usar el público en general, sin necesidad de que concurra ningún tipo de autorización. Existen tres tipos de datos en el dominio público:

1. La información no protegida por derechos de propiedad intelectual o *copyright*.
2. La información susceptible de quedar sujeta a protección, de conformidad con algún tipo de régimen de propiedad intelectual, pero que, de modo contractual, no queda protegida, como puede ser cuando se transfiera de forma abierta, permitiéndose su uso en abierto, sin reservas de derechos. Esto es relativamente frecuente en las colecciones de datos científicos.
3. La información que se pone a disposición de terceros haciendo uso de regímenes o excepciones para ello. Es decir, la que, en principio, podría tratarse de información protegida y que no lo está.

Existe toda una serie de razones lo suficientemente importantes como para promover el *open access* y el uso de los datos que genera el Gobierno, sin que ello tenga un coste para el público (Arzberger *et al.*, 2004). En efecto, el Gobierno no necesita incentivar al creador de información, habida cuenta de que el contribuyente ya ha pagado por ello en sus impuestos (Uhler y Reichman, 2003).

De alguna manera, limitar la diseminación abierta y el uso público de los datos e informaciones no solo puede comprometer la transparencia de los acuerdos de gobernanza y los valores democráticos, sino, también, los beneficios que la nación pudiera obtener de la diseminación ilimitada de todos estos datos.

La normativa estadounidense pone unos límites a los derechos de propiedad intelectual y excluye su aplicación a los datos generados por órganos o agencias gubernamentales. La consecuencia es una tendencia cada vez mayor a compartir los datos abiertamente y archivarlos en repositorios públicos. No obstante, esta no es la norma en ninguna otra parte, por lo que resulta muy difícil que Estados Unidos haga un uso similar de datos extranjeros, puesto que en ningún otro sitio el concepto dominio público es tan amplio como allí.

Aunque la norma general es que los datos de las agencias gubernamentales pasen a formar parte del dominio público, algunos factores pueden limitar el acceso libre, gratuito e indiscriminado a esos datos y a esa información. Estos factores limitativos se encuadran básicamente en tres categorías: problemas de seguridad nacional, la necesidad de proteger los datos personales de las personas que participen en investigación y la necesidad de respetar la información confidencial.

Existe, además, una cuarta limitación: el acceso a los datos no tiene que ser necesariamente gratuito.³⁷ Acceso abierto no es sinónimo de gratuito o sin coste, por lo que para acceder a dichos datos dicho coste puede variar.

Uhlir y Reichman (2003) concluyen que exigir costes superiores a los costes marginales puede crear sustanciales barreras de acceso, en particular en la investigación académica. Esta regulación difiere, prosiguen los autores, de la de la mayoría de los países en los que las agencias gubernamentales o cuasi-gubernamentales pueden explotar los datos obtenidos, y donde también son susceptibles de protección con arreglo a los derechos de propiedad intelectual.

Aunque, en la práctica, una de las mayores barreras de acceso es que las agencias no diseminan ni preservan a largo plazo los datos ni la información generados por el Gobierno. En última instancia, el problema respecto a este tipo de datos de dominio público es encontrarlos, es decir, poder acceder a ellos.

De hecho, ya en 2003, se señalaba la importancia de que fueran la National Science Foundation y otros organismos gubernamentales los que financiasen los proyectos relacionados con las ciberinfraestructuras (Atkins, 2003). El riesgo de no establecer

³⁷ Estas limitaciones tienen su origen en la Freedom Of Information Act (FOIA), que puede consultarse en la siguiente página: <http://www.dhs.gov/foia-exemptions>.

unas normas de creación y desarrollo de las ciberinfraestructuras puede suponer que las que se creen no sean compatibles o no puedan interactuar, etc.

Una quinta limitación que opera es la prohibición que afecta al Gobierno y que le impide competir con el sector privado a la hora de proporcionar información (tanto en términos de productos como en términos de servicios). Esta política limita, sustancialmente, el tipo de información que el Gobierno puede producir y divulgar (Uhlir y Reichman, 2003).

Por último, el Gobierno no puede divulgar los datos a los que tenga acceso y que sean propiedad del sector privado, que se han puestos a disposición del Gobierno para que este los use, salvo que exista algún tipo de exención (Uhlir y Reichman, 2003).

B. Los datos generados por instituciones académicas e instituciones sin ánimo de lucro

Los datos y la información obtenidos de la investigación financiada por organizaciones filantrópicas no acceden, de forma directa y automática, al dominio público. A diferencia de las anteriores, debe hacerse activamente algo para que esto suceda. En caso contrario, por defecto, se les atribuirán derechos de propiedad intelectual o *copyright*.

Una recomendación general del National Research Council (Consejo Nacional de Investigación de Estados Unidos) hizo hincapié en la necesidad de poner los datos y la información resultante de estas investigaciones, en la medida de lo posible, en el dominio público y, por ende, en no aplicar restricciones de ningún tipo.

El hecho de que estos datos entren en el dominio público, sin embargo, está ligado a los requerimientos contractuales de las agencias a la hora de dar subvenciones, así como a una serie de normas aplicables a la ciencia. La aspiración final de dichas normas es implementar un *open access* completo en lo que respecta a los datos científicos, así como compartir los resultados de la investigación con la finalidad de alimentar nuevos proyectos.

Esta política de *open access* está definida en varios documentos de Gobierno de Estados Unidos. De hecho, el National Research Council lo reporta en los siguientes términos:

Los datos e información derivados de la investigación subvencionada con fondos públicos deben ponerse a disposición de los que quieran hacer uso de ellos con el mínimo de restricciones posibles, de forma no discriminatoria, sin imponerles un coste superior a la reproducción y distribución. Estados Unidos tiene, y ha tenido, desde los años cincuenta, elevadas inversiones en promover la investigación, basada en dos presupuestos básicos: el no-coste de las transferencias y de la reutilización de la información (Pavitt, 2001).

El fundamento está en la interpretación de que la inversión en investigación básica es necesaria para proporcionar las competencias y el conocimiento necesarios para mantener la competitividad y el nivel de desarrollo. Un gran número de agencias de Estados Unidos están promoviendo estos principios con la finalidad de que la investigación básica y los datos ligados a la misma, subvencionados con fondos públicos, se conviertan así en *inputs* en el dominio público.

De este modo, los regímenes tradicionales de *copyright* han puesto de manifiesto que hechos, ideas, datos y descubrimientos forman parte del dominio público. Esto, además, tiene sentido teniendo en cuenta que no hay ninguna actividad creadora, sino tan solo observación de lo que ocurre en la naturaleza. No parece que tenga ningún tipo de justificación proteger este tipo de información con las normas de *copyright*.

A todo el conocimiento de las organizaciones y proyectos con subvenciones gubernamentales deben sumarse los datos que obran en poder de los científicos y que son compartidos con modelos contractuales. La finalidad de estas cláusulas contractuales no es otra que la de asegurarse de que los datos que recaben o generen las personas físicas o jurídicas que reciban la subvención se compartirán abiertamente con otros investigadores (sin perjuicio de que puedan imponerse periodos de uso exclusivo o embargos). En muchos casos, los datos se publican en archivos públicos después de su publicación.

No obstante, están proliferando prácticas en las que se incentiva la comercialización de estos datos para obtener nuevas formas de financiación, como se verá después. La ley Bayh-Dole de 1980 incentiva sobre todo a los académicos a proteger los resultados de su investigación en los casos relacionados con la investigación en el ámbito de la biomedicina, y esta ley se aplica a los resultados que sean patentables.

Existe también, en este ámbito, un entorno en el que se fomentan los intercambios de datos informales. Esta tendencia a compartir la información y los resultados es más frecuente en grandes proyectos que en pequeños proyectos. En efecto, en esos casos, los investigadores son más reticentes a poner los resultados obtenidos —y con más razón los procesos que han utilizado para alcanzar dichos resultados— a disposición del público en general. Aunque, sin lugar a dudas, esta forma de generar e intercambiar datos es clave (Perkmann y Walsh, 2009).

C. El sector privado

El sector privado es un claro consumidor de datos de dominio público y al mismo tiempo un gran contribuidor al desarrollo de los mismos (Uhlir y Reichman, 2003). De hecho, el sector privado es uno de los grandes productores de datos que

permiten el desarrollo de los datos en otros ámbitos como puede ser en el ámbito académico.

Como consecuencia de la mencionada ley Bayh-Dole, se han desarrollado las relaciones entre el mundo académico y el empresarial. La cada vez mayor interacción entre el mundo académico y el empresarial genera un conflicto entre las normas de

open access a las que estaba sujeta la investigación tradicional en las universidades frente a la tendencia que predomina en el sector privado de accesos restringidos y limitaciones en el uso de los datos, normalmente sujetos, contractualmente, a confidencialidad.

Este tema no tiene todavía una respuesta. Más adelante, junto con las cuestiones que se plantean en el marco del *copyright*, se analizarán en el capítulo relativo a la interoperabilidad legal.

3.2.1.3. Un análisis comparado entre la protección de las bases de datos en Europa y en Estados Unidos: conclusiones

Por lo tanto, en Europa se adopta un fuerte derecho de propiedad en colecciones no susceptibles de ser protegidas mediante los derechos de propiedad intelectual al adoptarse una protección de la estructura mediante la regulación de la protección de las bases de datos (Reichman y Uhler, 1999).

Como se ha visto hasta ahora, existen una serie de diferencias clave entre ambas regulaciones. En Estados Unidos, la protección de *copyright* abandona la protección del *sweat of the brow* y exige unos requisitos adicionales para la aplicación de la protección de las bases de datos. En efecto, se necesita un cierto grado de originalidad para poder proteger las bases de datos. De alguna manera, esto da un mayor acceso a las bases de datos y limita las restricciones que pueden imponerse.

Sin embargo, Europa recoge este principio del *sweat of the brow* y adopta una política normativa mucho más protectora en lo que respecta a la protección de las bases de datos. Con estas restricciones, persiguen proteger el mercado europeo de las bases de datos y conseguir un mayor desarrollo de las mismas en la Unión. Algunos autores ponen de manifiesto que el objetivo no se ha alcanzado.

Una de las particularidades de la directiva europea es la creación del ya mencionado derecho *sui generis*, que no tiene parangón en otros derechos. Desde que se adoptara en el derecho europeo, Estados Unidos se ha planteado varias veces la posibilidad de recogerlo en su normativa pero, no lo ha hecho hasta ahora.

Otra de las diferencias fundamentales es que el Estados Unidos existe la excepción denominada *fair use*, mientras que en Europa se recogen una serie de limitaciones y excepciones al uso.

No obstante, existe una condición previa a la existencia de ambos sistemas que ha cristalizado de forma diferente.

La Convención de Berna, en su revisión de 1967, exige que cualesquiera limitaciones que se apliquen a las normas de derecho de propiedad intelectual satisfagan el test de los tres pasos (Max Planck, s. f.). Se establece que podrá permitirse la reproducción de dichos trabajos siempre que se regule en los distintos Estados: en algunos casos concretos, siempre que dicha reproducción no entre en conflicto con la explotación normal del trabajo, y no perjudique de forma no razonable los intereses legítimos del autor.

Por consiguiente, se va a proceder a analizar la forma de implementarlo en Europa, es decir, sus limitaciones y excepciones. Estas tienen su origen en la forma continental de regular en Europa que las recoge de forma extensa y exhaustiva, pues lo que no esté recogido y descrito en normas genera inseguridad jurídica en la aplicación. Las excepciones están recogidas y listadas en el artículo 5 de la directiva (Parlamento Europeo y Consejo de la Unión Europea, 2001) que establece:

1. Los actos de reproducción provisional a que se refiere el artículo 2 [los derechos de reproducción], que sean transitorios o accesorios y formen parte integrante y esencial de un proceso tecnológico y cuya única finalidad consista en facilitar:
 - a. una transmisión en una red entre terceras partes por un intermediario, o
 - b. una utilización lícita de una obra o prestación protegidas, y que no tengan por sí mismos una significación económica independiente, estarán exentos del derecho de reproducción contemplado en el artículo 2.
2. Los Estados miembros podrán establecer excepciones o limitaciones al derecho de reproducción contemplado en el artículo 2 en los siguientes casos:
 - a. en relación con reproducciones sobre papel u otro soporte similar en las que se utilice una técnica fotográfica de cualquier tipo u otro proceso con efectos similares, a excepción de las partituras, siempre que los titulares de los derechos reciban una compensación equitativa;

- b. en relación con reproducciones en cualquier soporte efectuadas por una persona física para uso privado y sin fines directa o indirectamente comerciales, siempre que los titulares de los derechos reciban una compensación equitativa, teniendo en cuenta si se aplican o no a la obra o prestación de que se trate las medidas tecnológicas contempladas en el artículo 6;
 - c. en relación con actos específicos de reproducción efectuados por bibliotecas, centros de enseñanza o museos accesibles al público, o por archivos, que no tengan intención de obtener un beneficio económico o comercial directo o indirecto;
 - d. cuando se trate de grabaciones efímeras de obras, realizadas por organismos de radiodifusión por sus propios medios y para sus propias emisiones; podrá autorizarse la conservación de estas grabaciones en archivos oficiales, a causa de su carácter documental excepcional;
 - e. en relación con reproducciones de radiodifusiones efectuadas por instituciones sociales que no persigan fines comerciales, como hospitales o prisiones, a condición de que los titulares de los derechos reciban una compensación equitativa.
3. Los Estados miembros *podrán establecer excepciones o limitaciones* a los derechos a que se refieren los artículos 2 [derechos de reproducción] y 3 [derechos de comunicación al público] en los siguientes casos:
- a. cuando el *uso* tenga únicamente por objeto la *ilustración con fines educativos o de investigación científica*, siempre que, salvo en los casos en que resulte imposible, se indique la fuente, con inclusión del nombre del autor, y en la medida en que esté justificado por la finalidad no comercial perseguida;
 - b. cuando el *uso* se realice en beneficio de *personas con minusvalías*, guarde una relación directa con la minusvalía y no tenga un carácter comercial, en la medida en que lo exija la minusvalía considerada;
 - c. cuando la *prensa* reproduzca o se quiera comunicar o poner a disposición del público artículos publicados sobre temas de *actualidad económica, política o religiosa*, o emisiones de obras o prestaciones del mismo carácter, en los casos en que dicho uso no esté reservado de manera expresa, y siempre que se indique la fuente, incluido el nombre del autor, o bien cuando el uso de obras o prestaciones guarde conexión con la información sobre acontecimientos de actualidad, en la medida en que esté justificado por la finalidad informativa y siempre que, salvo en los casos en

que resulte imposible, se indique la fuente, con inclusión del nombre del autor;

- d. cuando se trate de *citas con fines de crítica o reseña*, siempre y cuando estas se refieran a una *obra* o prestación que se haya puesto ya *legalmente a disposición del público*, se indique, salvo en los casos en que resulte imposible, la fuente, con inclusión del nombre del autor, y se haga buen uso de ellas, y en la medida en que lo exija el objetivo específico perseguido;
- e. cuando el uso se realice con *fines de seguridad pública* o para garantizar el *correcto desarrollo de procedimientos administrativos*, parlamentarios o judiciales, o para asegurar una cobertura adecuada de dichos procedimientos;
- f. cuando se trate de *discursos políticos y de extractos de conferencias públicas* u obras o prestaciones protegidas similares en la medida en que lo justifique la finalidad informativa y siempre que, salvo en los casos en que resulte imposible, se indique la fuente, con inclusión del nombre del autor;
- g. cuando el *uso* se realice durante *celebraciones religiosas o celebraciones oficiales* organizadas por una autoridad pública;
- h. cuando se usen *obras*, tales como obras de arquitectura o escultura, realizadas para estar situadas de forma permanente en *lugares públicos*;
- i. cuando se trate de una inclusión incidental de una obra o prestación en otro material;
- j. cuando el uso tenga la finalidad de anunciar la *exposición pública* o la venta de *obras de arte*, en la medida en que resulte necesaria para promocionar el acto, con exclusión de cualquier otro uso comercial;
- k. cuando el uso se realice a efectos de *caricatura, parodia o pastiche*;
- l. cuando se use en relación con la *demostración o reparación de equipos*;
- m. cuando se use una obra de arte en forma de edificio o dibujo o plano de un edificio con la *intención de reconstruir dicho edificio*;
- n. cuando el uso consista en la *comunicación a personas concretas* del público o la puesta a su disposición, a efectos de *investigación* o de *estudio personal*, a través de terminales especializados instalados en los locales de los establecimientos mencionados en la letra c) del apartado 2, de obras y prestaciones que figuran en sus colecciones y que no son objeto de condiciones de adquisición o de licencia;

- o. cuando el uso se realice en otros casos de importancia menor en que ya se prevean *excepciones o limitaciones en el Derecho nacional*, siempre se refieran únicamente a usos analógicos y que no afecten a la libre circulación de bienes y servicios en el interior de la Comunidad, sin perjuicio de las otras excepciones y limitaciones previstas en el presente artículo.
4. Cuando los Estados miembros puedan establecer excepciones o limitaciones al derecho de reproducción en virtud de los apartados 2 y 3, podrán igualmente establecer excepciones o limitaciones al derecho de distribución previsto en el artículo 4, siempre que lo justifique la finalidad del acto de reproducción autorizado.
5. Las excepciones y limitaciones contempladas en los apartados 1, 2, 3 y 4 únicamente se aplicarán en determinados casos concretos que no entren en conflicto con la explotación normal de la obra o prestación y no perjudiquen injustificadamente los intereses legítimos del titular del derecho.

A diferencia de lo que ocurre en otros aspectos de la directiva, el listado de limitaciones y excepciones no resulta de un análisis comparado de normativas, ni tampoco fueron formuladas en función de ningún objetivo concreto (Cook, 2012).

Una categorización que plantea Gervais (2008) de las excepciones y limitaciones podría ser la siguiente:

CATEGORIZATION	CATEGORIES	INTERNAL BALANCE	EXTERNAL NORMS
By type of user	Limited ability users	Braille copies	Non-discrimination
	Consumers	Private sphere/ difficult enforcement	Privacy, consumer protection
	Government		Education, culture, information (national security)
	Institutional ⁸⁰		Education, culture, information
	All	Reuse by quotation	Information, free expression
By type of use	Consumptive	Private sphere/ difficult enforcement	Privacy, consumer protection, education, information, culture
	Creative/transformativa	Limit right to prohibit when beyond need; public interest balance	Free expression, culture, information
	Informational	Public interest balance	Information, free expression
By type of country	Developing country		Right to development; education
By type of author	Governmental works	No incentive needed	Information
By type of work	Computer software	Public interest function does not require prohibition of reverse engineering ⁸¹	Competition
	Printed publications	Access does not interfere with copyright's function	Education, information

Figura 25. Posibles principios para una conceptualización de limitaciones y excepciones. Fuente: Gervais, 2008.

Dentro de esta normativa un tema muy relevante que llama la atención es analizar la excepción relativa a la gestión de *text and data mining* (minería de texto y datos) que recoge dicho borrador de directiva.

Procede hacer alguna reflexión previa sobre las cuestiones que suscita la minería de texto y datos (Korn *et al.*, 2007). En efecto, estos autores señalan que una de las condiciones para que exista la reutilización de los datos es el acceso a los mismos. No cabe duda de que una de las formas de acceso es la minería de texto y datos. Por lo tanto, los datos derivados llevan aparejados una complejidad de gestión en lo que respecta a los derechos de propiedad intelectual.

En lo que respecta a la definición de *fair use*, se va a proceder a analizar la función que desempeña en el ámbito del *copyright* (Burk y Cohen, 2001). Y es que esta figura desempeña una variedad de funciones, todas ellas relacionadas dentro del marco de las regulaciones del *copyright*. Una de las bases para la creación del *fair use* es la bondad de compartir, de intercambiar (*sharing is good*) (Madison, 2004). Sin embargo, en el marco de los derechos de propiedad intelectual, este sentido de que compartir es bueno suscita mucha controversia.

Por definición, prosigue Madison (2004), esta herramienta se configura como una manera de permitir el intercambio de datos o información no autorizados de trabajos que tengan un contenido creativo.

Esta especie de trampa en el sistema de protección ha suscitado una multiplicidad de cuestiones que se han ido analizando a lo largo de los años para construir lo que hoy es el *fair use*. A fin de cuentas, esta herramienta es una forma de limitar el monopolio derivado de la aplicación del *copyright* (Leval, 1990). Aunque el reconocimiento de la protección del *copyright* supone el reconocimiento de que es una forma de proteger la creatividad de los autores, de alguna manera todas las obras tienen su origen en otras, es decir, son obras derivadas. Por lo tanto, una protección de *copyright* que impidiera cualquier referencia a otras obras ahogaría los nuevos procesos creativos.

El uso legítimo o uso razonable (Fair use, s. f.) es un criterio jurisprudencial desarrollado en el sistema del derecho anglosajón (o *common law*). Permite un uso limitado de material protegido sin que sea necesario permiso del dueño de dichos derechos, por ejemplo: para uso académico o informativo.

Permite la cita o incorporación, legal y no licenciada, de material protegido en un trabajo de otro autor, bajo el requerimiento de cumplir cuatro condiciones.

1. El propósito y el carácter del uso en disputa.
2. La naturaleza de la obra protegida con derechos de autor.
3. La importancia de la parte utilizada en relación con la obra en total.
4. El efecto de dicho uso en el mercado sobre el valor de la obra protegida con derechos de autor.

El sistema del *fair use* se está extendiendo a más países que Estados Unidos, lo que lleva a algún autor a preguntarse si existe una tendencia que vaya a convertir el *fair use* en un concepto de aplicación más global (Decherney, 2014). De hecho, algunos países de la Unión Europea consideran el sistema europeo demasiado inflexible y están analizando formas de acercarse a la regulación del *fair use* (Cook, 2012).

Sin lugar a dudas, existe hoy una necesidad de flexibilizar las normas de derecho de propiedad intelectual en la Unión Europea. La necesidad de más apertura (*openness*) en el derecho de propiedad intelectual es casi indiscutible en la sociedad de la información de cambios crecientes, dinámicos e imprevisibles (Hugenholtz y Senftleben, 2011). De hecho, parece que el derecho de propiedad intelectual ha perdido gran parte de su flexibilidad en las jurisdicciones de derecho civil en Europa, durante el último siglo. No obstante, eso no quiere decir necesariamente que deba introducirse el *fair use*.

Como consecuencia de ello Hugenholtz y Senftleben (2011) recomiendan introducir medidas de flexibilidad que convivan con la estructura existente de excepciones y limitaciones y que, por lo tanto, combinen las ventajas de potenciar la flexibilidad con la seguridad jurídica y la neutralidad tecnológica.

Parece que, aunque la introducción del *fair use* en Europa es una suerte de tabú (Hugenholtz y Senftleben, 2011), sí se está extendiendo la idea de que es necesario flexibilizar las interpretaciones de las excepciones y limitaciones dando cabida a las cada vez más frecuentes situaciones nuevas que se generan en el ámbito de la sociedad del conocimiento.

De hecho, un informe elaborado por el Committee on Legal Affairs (2015) de la Unión Europea señala la necesidad de analizar el creciente impacto de las diferencias que existen entre Estados miembros a la hora de implementar las excepciones por la incertidumbre legal que supone y por los efectos negativos que produce en el mercado común digital. Aunque se pide flexibilizar, todavía se está lejos del *fair use*.

No obstante, algunos autores ya hablan de *fair use* o, cuando menos, de acercamiento al mismo. En efecto, la flexibilización de los sistemas se va acercando al concepto de *fair use* (Malcolm, 2015).

3.2.1.4. Una breve reflexión sobre los datos de biodiversidad

Las restricciones que se van imponiendo a los datos de biodiversidad son cada vez mayores, sobre todo en lo que respecta a su protección o, lo que es lo mismo, la tendencia es limitar cada vez más su protección y fomentar que se compartan. Esto es así tanto porque la mayor parte de los datos que se recogen se hacen en el marco de investigaciones subvencionadas como porque los datos y observaciones de biodiversidad son reproducciones de la realidad.

Un ejemplo de dato podría ser una imagen obtenida mediante los sistemas de observación actuales. Algunos autores como Egloff *et al.* (2017) cuestionan abiertamente que una imagen sea susceptible de protección con arreglo a los derechos de propiedad intelectual. De hecho, reclaman que los dibujos o fotografías reproducen una realidad observable y que, si reúnen determinadas condiciones, pueden no ser susceptibles de ningún tipo de protección.

En lo que afecta a los mapas, por ejemplo, se plantean otras cuestiones. Si bien es cierto que los mapas no pueden protegerse, a la hora de elaborarlos es importante usar herramientas no protegidas por derechos de autor o patentes.

Muchos autores hablan ya de la necesidad de analizar qué pesa más: la protección de los datos o bien el interés general. De hecho, la protección de los datos contenidos en las bases de datos de biodiversidad podría servir para paralizar todo el

intercambio de datos, impidiendo afrontar de forma global las cuestiones que la misma plantea, por ejemplo: la modelización, la reacción a determinados desastres, etc.

Aunque el problema jurídico está en revisión, lo cierto es que, en muchas de las bases de datos de biodiversidad, se pone como condición previa el intercambio de los datos.

3.3. Las tendencias *open* (*open data* y *open access* o acceso abierto): el nuevo marco para el intercambio de datos

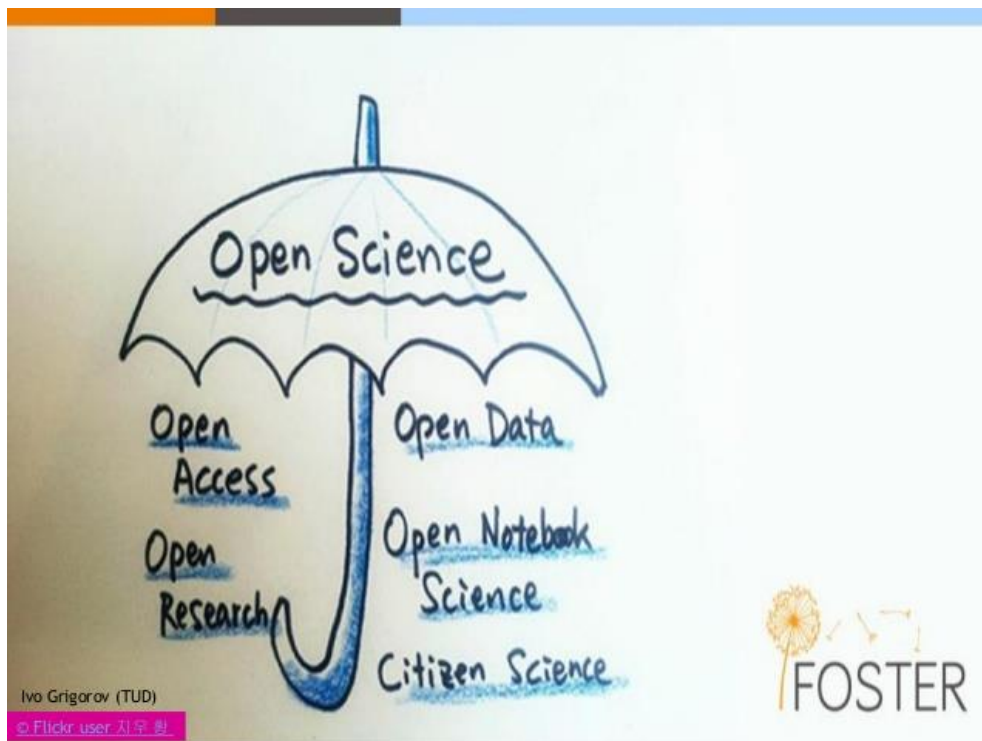


Figura 26. El paraguas de la Open Science. Fuente Grigorov³⁸.

El *open science* para hacerse realidad o cristalizar necesita de las tendencias *open* (*open data* y *open access* o acceso abierto se conocen como OA, por sus siglas en inglés (en adelante, se usarán de forma indistinta). Como ya se ha comentado, los derechos de propiedad intelectual o *copyright* nacieron como una manera de garantizar la innovación, asegurando la publicación de los datos. Sin embargo, pueden haberse convertido hoy en un freno frente al fenómeno de la digitalización.

³⁸ Recuperado de:

https://www.google.de/search?q=ivo+grigorov+open+science&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiljb7AoZTaAhVHVRQKHUt3CsAQ_AUICygC&biw=1384&bih=627#imgsrc=2_xN63uO-AsUkM: (Consultado por última vez el 14 de marzo de 2018).

Empieza a surgir un fuerte movimiento político y social que cuestiona la necesidad de adecuar los derechos de propiedad intelectual o *copyright* a la nueva realidad. Se centran en una nueva tendencia: el «acceso abierto». Este, encuentra su fundamento en una necesidad de soluciones más global, en la que empiezan a aparecer nuevas herramientas y enfoques.

3.3.1. Las distintas modalidades de open

3.3.1.1. Open access o acceso abierto (OA)

La configuración del concepto de *open access* empieza a fraguarse en el contexto de lo que se conoce como las BBB (3Bs) del *open access*. Las BBB se corresponden con las iniciales de Budapest (2002), Bethesda (2003) y Berlín (2003), que fueron los lugares donde se celebraron las reuniones de las tres declaraciones que sirvieron de base para la construcción conceptual del *open access*.

La *Open Initiative de Budapest* (BOAI) (Open Society Institute, 2001) de febrero de 2002 manifiesta:

Por acceso abierto a la literatura (científica), entendemos su disponibilidad gratuita en Internet, para que cualquier usuario la pueda leer, descargar, copiar, distribuir o imprimir, con la posibilidad de buscar o enlazar al texto completo, recoger los artículos para su indexación, pasarlos como datos para software o utilizarlos para cualquier otro propósito legítimo, sin más barreras financieras, legales o técnicas que aquellas que supongan acceder a Internet. El único límite a la reproducción y distribución de los artículos publicados y la única función del copyright en este marco, tiene que ser garantizar a los autores el control sobre la integridad de su trabajo y el derecho a ser reconocido y citado (Barrionuevo, 2009).

La *Declaración de Bethesda sobre Publicación de Acceso Abierto* de abril de 2003 va un poco más allá y añade:

Para que un trabajo tenga la condición de acceso abierto, debe reunir dos condiciones:

1. El(los) autor(es) y el(los) propietario(s) de los derechos de propiedad intelectual concede(n) a todos los usuarios un derecho de acceso a los datos libre, irrevocable, universal y perpetuo y una licencia para copiar, utilizar, distribuir, transmitir y presentar el trabajo públicamente y hacer y distribuir obras derivadas, en cualquier soporte digital, para cualquier finalidad responsable, siempre que se realice una adecuada atribución de autoría, así como el derecho de hacer una pequeña cantidad de copias impresas para su uso personal.

2. Una versión completa de la obra y todos los materiales adicionales, en la que se incluya una copia de los anteriormente citados permisos, en un formato electrónico estándar apropiado, deberá depositarse de forma inmediata a la publicación inicial en al menos un repositorio en línea apoyado por una institución académica, una sociedad académica, una agencia gubernamental, o cualquier otra organización debidamente establecida que persiga posibilitar el acceso abierto, la distribución sin restricciones, la interoperabilidad y el archivo a largo plazo [...].»

A su vez, la *Declaración de Berlín sobre el Acceso Abierto al Conocimiento en Ciencias y Humanidades*, de octubre de 2003, afirma:

Nuestra misión de diseminar el conocimiento será incompleta si la información no es puesta a disposición de la sociedad de manera expedita y amplia. Es necesario apoyar nuevas posibilidades de diseminación del conocimiento, no solo a través de la manera clásica, sino también utilizando el paradigma de acceso abierto por medio de Internet. Definimos el acceso abierto como una amplia fuente de conocimiento humano y patrimonio cultural aprobada por la comunidad científica.

Para que pueda alcanzarse la visión de una representación del conocimiento global y accesible, la Web del futuro tiene que ser sostenible, interactiva y transparente. El contenido y las herramientas de *software* deben ser libremente accesibles y compatibles.

En estas tres declaraciones y convenciones tiene su origen la definición conocida como BBB del *open access*. También por esa razón, el año 2003 se conoce como el año del acceso abierto (Willinsky, 2006).

Existen otras muchas definiciones de acceso abierto, por ejemplo Harnard (2006a), considerado uno de los fundadores de la iniciativa, afirma: «mi definición de *open access* es la misma que la de la Declaración de Budapest: dar acceso sin restricciones y a través de Internet a los textos completos de la literatura científica ya revisada. Aunque esta definición omite dos adjetivos importantes, inmediato y permanente». El autor, en la misma obra, va un paso más allá y pone de manifiesto que, de no existir, el impacto de los artículos se ve reducido y, con él, la diseminación del conocimiento.

En la revolución del *open access* han existido dos etapas bien diferenciadas: la primera, el paso de las publicaciones en papel a la digitalización de las mismas y, la segunda, el acceso sin restricciones a este tipo de publicaciones (Björk *et al.*, 2010).

En la primera de las fases, la digitalización de las publicaciones es la que, por primera vez, plantea el acceso sin limitaciones al conocimiento en la red (MacGregor y Stranack, s. f.). Aunque al principio con reservas, es creciente la

preferencia de los trabajos académicos por la red como el método de diseminación. Por ejemplo, el proyecto Public Knowledge nace como un modesto proyecto en la Universidad de Columbia con la finalidad de reducir las barreras para promover un mayor acceso al conocimiento. De hecho, realiza un análisis de los condicionantes que existen (Willinsky, 1999).

La segunda etapa es el acceso sin restricciones en sí mismo, el acceso abierto. No existe un único concepto, una única definición de *open access*, aunque ha servido de punto de partida, la definición de Suber (2012). Así, la literatura *open access* es literatura digital, accesible vía Internet, sin coste y lo más abierta posible tomando como referencia los derechos de autor o *copyright*. Es decir, se enfoca en hacer la información lo más reutilizable posible, que prácticamente no existan restricciones en su uso.

El *open access* elimina las barreras de precios (suscripciones, costes de licencia, costes de consulta) y las barreras legales (restricciones de *copyright* o derechos de autor y de licencias). Así, los puntos más importantes de la definición son la libre disposición y la inexistencia de restricciones de acceso. Estos dos términos captan ambos elementos.

Existe flexibilidad en lo que respecta a las barreras legales a eliminar (Suber, 2012). Por ejemplo, algunos proveedores de *open access* permiten usos comerciales, mientras que otros no los permiten. Algunos permiten trabajos derivados y otros, no, pero la mayoría de definiciones públicas de *open access* confluyen. No es suficiente limitarse a suprimir las barreras de los costes o restringir los usos permitidos a los derivados del *fair use* en Estados Unidos o *fair dealing* en Reino Unido.

open access es el término usado para describir un modelo radicalmente nuevo para diseminar las publicaciones científicas de investigación. Está sirviendo para sustituir, gradualmente, los métodos tradicionales de pago de suscripciones a revistas y de accesos restringidos a solo lectores (Björk et al., 2013).

Aunque las BBB han servido de base para la construcción del concepto de *open access*, el concepto se ha demorado más tiempo en arraigar y en trasladarse a las políticas y exigencias de proyectos, publicaciones e instituciones.

La existencia del acceso abierto ha suscitado toda una serie de análisis para evaluar si aumenta el número de citas de los artículos publicados usando este sistema de acceso abierto (Craig et al., 2007).

De la definición, extraída de las declaraciones procede señalar varios puntos clave (Bailey, 2007). En primer lugar, las obras de acceso abierto están disponibles gratuitamente. En segundo lugar, están en línea, lo que normalmente significa que son documentos digitales disponibles en Internet. En tercer lugar, son trabajos

académicos y, por lo tanto, están excluidos las novelas románticas, las revistas populares, los libros de autoayuda y similares.

En cuarto lugar, los autores de estas obras no reciben pago alguno a cambio de sus esfuerzos. En quinto lugar, dado que la mayoría (pero no todos) los autores de artículos de revistas revisados por pares no son remunerados y que tales trabajos son académicos, estos artículos se identifican como el tipo principal de material de acceso abierto. En sexto lugar, hay un número extraordinario de usos permitidos para materiales de acceso abierto. Los usuarios pueden copiar y distribuir las obras de acceso abierto sin restricciones, sin perjuicio de los requisitos de correcta atribución del autor y la garantía de la integridad de la obra. En séptimo y último lugar, existen dos estrategias clave de acceso abierto: las revistas de autoarchivado y las de acceso abierto.

Los que defienden el *open access* están convencidos de que la comunicación científica puede mejorarse si, a través de Internet, se hace posible el acceso a todo el contenido académico sin restricciones y sin coste (Abadal, 2013).

El acceso abierto promueve los principios de *open science*, que elimina las barreras en relación con la participación en la investigación y en la formación. Contribuye, además, al bien común (Borgman, 2010). *Open* tiene una connotación de mínimas restricciones de uso y esto no equivale necesariamente a que el uso sea gratuito. El contenido tiene que estar disponible, sin que sea necesario que lo esté de forma gratuita (Borgman, 2010).

El *open access* tiene distintas modalidades. En la actualidad, se han diferenciado, básicamente, dos tipos: la ruta verde (*Green open access*) y la ruta dorada (*Golden open access*). Esta terminología, como tal, tiene su origen en el año 2004, aunque los conceptos son anteriores (Björk *et al.*, 2013).

Son dos soportes o vehículos distintos para ofrecer el *open access*. En la ruta o acceso abierto verde, los propietarios de una publicación la ponen a disposición del público directamente en un repositorio. El riesgo de que este vehículo se use es tremendamente dependiente de que los investigadores carguen los documentos en el repositorio, aunque es verdad que algunas instituciones han comenzado a exigirselo a sus investigadores. De hecho, se asimila la ruta verde con el hecho de que el autor lo ponga a disposición. Sin embargo, Björk *et al.* (2013) incluyen en esta ruta todas las copias accesibles de los artículos, incluidas las distintas versiones de dichos artículos, que existan en páginas web distintas de las del editor original.

En la ruta dorada, en cambio, se ofrece el *open access* a través de revistas, de publicaciones de acceso abierto. En este caso, los artículos y publicaciones son revisadas por pares. Es la forma más sostenible a largo plazo.

No cabe duda de que existe un tremendo impacto de ambas rutas (Harnard *et al.*, 2004). De hecho, existe un debate sobre cuál de las dos opciones debe prevalecer (Abadal, 2013). Algunos autores ponen de manifiesto que ha quedado claramente demostrado que diseminar los resultados de investigación mediante el acceso abierto resulta más eficiente en costes que las suscripciones. En un mundo de acceso abierto parece que los beneficios netos de la ruta dorada superan a los de la verde. Por el momento, aunque no se haya llegado aún a ese punto, la mejor manera de desarrollar el acceso abierto es recurrir a la vía verde (Houghton y Swan, 2013).

Además, existen dos modalidades de *open access*: *open access* libre, que supone un acceso sin costes y la posibilidad de hacer alguno o todos los usos, y un acceso *open access* gratis, es decir, un acceso sin costes (Bailey, 2010).

El *open access* es, por lo tanto, la base de todo el planteamiento de intercambio de datos o *data sharing* y es necesario para la conservación de la biodiversidad (Fonseca y Benson, 2003).

Existen ya ejemplos de organizaciones o fundaciones privadas que prohíben a sus investigadores publicar datos en revistas que no sigan las políticas de *open access*, pues entienden que están limitando el desarrollo de la ciencia. El último ejemplo es la prohibición de la Fundación Gates a sus científicos de publicar en revistas como *Nature*, *Science*, *The New England...*, precisamente por las restricciones a la reutilización de los datos a las que quedan sujetos los artículos de dichas publicaciones.

Aparte de la importancia de aplicar el acceso abierto a los datos de investigación (Wilbanks, 2006), una parte importante del debate sobre acceso abierto se ha centrado en el principio de acceso por parte de los científicos y las consecuencias económicas de este cambio.

No obstante, existe un segundo principio a considerar, tal como pone de relieve Wilbanks (2006), y es el pleno funcionamiento de los nuevos enfoques tecnológicos como pueden ser la minería de datos, los filtros colaborativos y la indexación semántica, que no están generando nuevas fuentes públicas. Y es que, a pesar de los grandes avances del acceso abierto, la mayor parte de la investigación académica no está disponible, ni para estudios ni para que la procese ningún *software*.

No cabe duda de que la apertura en la ciencia es crítica y que es necesario articular modelos de compartir los datos. A fin de cuentas, hoy el dato es la base de la economía y del desarrollo, pues afecta desde al autobús que pasa hasta las investigaciones de una multiplicidad de enfermedades o el genoma humano. Parece difícil avanzar sin estos presupuestos de compartir (Wilbanks, 2011).

Aparecen los FAIR Data Principles, sobre los que se volverá a continuación. *FAIR* es un conjunto de principios rectores con la finalidad de hacer los datos *findable* (fáciles de encontrar), *accessible* (accesibles), *interoperables* (interoperables) y *reusable* (reutilizables). Muchas instituciones han optado por la aplicación de algunos de ellos. Además, el Programa Horizonte 2020 los ha adoptado como base en sus políticas. De igual manera, en 2016, el G20 endosó los FAIR Principles. Sin embargo, el consenso acerca de los mismos es discutible.

3.3.1.2. Open data

Un reflejo del *open* en los datos es el llamado *open data*. Este término fue acuñado por Peter Murray-Rust con la finalidad de distinguir entre los datos comprensibles para un ordenador de otras formas de acceso abierto (Borgman, 2010). Es un término muy específico y hace referencia a dos tipos de apertura:

1. *Apertura legal*. Los datos están legalmente abiertos, lo que en la práctica significa que están publicados bajo una licencia abierta y que las condiciones de reutilización están limitadas a la obligación de citar.
2. *Apertura técnica*. Los datos están técnicamente abiertos, es decir, las máquinas pueden leer los ficheros y estos están accesibles para todo el mundo, además de que el formato de los ficheros y sus contenidos no están restringidos a ningún *software* propietario en particular (Bernal y Román Molina, 2016).

Si el concepto de red ha supuesto una revolución en lo que respecta al acceso a los datos de alguna manera, la interoperabilidad es una de las condiciones para la existencia del *open data*, pues los datos de forma aislada no sirven para nada (Rizzo *et al.*, 2009).

De acuerdo con su fundador Tim Berners-Lee (2000), la finalidad de la red es dar acceso a la información. En su origen, está el *open data* (Gurstein, 2011). En esta misma línea, la ciencia se basa en los datos: su colección, análisis, publicación, reanálisis, crítica y reutilización (Molloy, 2011).

El sistema actual de publicación científica trabaja en contra de la máxima difusión de los datos científicos. Las barreras incluyen la imposibilidad de acceder a los datos, las restricciones de uso aplicadas por los editores o los proveedores de datos y la publicación de datos difíciles de reutilizar porque están, por ejemplo, mal anotados u «ocultos» en tablas no modificables como documentos PDF. Existe una renuencia cultural a publicar datos abiertamente, por múltiples razones. Sin lugar a dudas, los investigadores temen abrir, compartir los datos, por miedo a perder el control sobre su uso. Carecen de incentivos para ello y no quieren perder el crédito. Todas estas cuestiones no contribuyen a la apertura de los datos.

El concepto de *open data* se está construyendo mediante el proceso de definir de qué manera los datos científicos pueden publicarse y reutilizarse sin coste y sin barreras. Los científicos, por lo general, consideran que los datos publicados pertenecen a la comunidad científica, pero muchos editores reclaman *copyright* sobre los datos y prohíben su reutilización sin autorización. Esto es un impedimento para el desarrollo del conocimiento en la era digital (Murray-Rust, 2008).

Este concepto es la base de la construcción de un concepto de ciencia mucho más globalizado (Uhlir y Schröder, 2007); de hecho, la construcción del concepto de ciencia global está en un punto crítico.

En lo que respecta a los datos de biodiversidad, el *open data* es, sin duda, clave. Abordar un desafío mundial como la pérdida de biodiversidad requiere ensamblar productos de información globales a través de múltiples escalas espaciales y temporales. La teledetección por satélite es especialmente útil para generar registros de observación consistentes de los factores clave del cambio de la biodiversidad (es decir, la cobertura de la tierra y la dinámica del uso de la tierra, las variables climáticas y las condiciones de la superficie del mar) desde un nivel local hasta uno global (Hansen y Loveland, 2012).

Why make data available?

"It was **never** acceptable to publish papers without making data available."

- Ewan Birney

#OpenData
#OpenScience



Original image via doi:10.1038/461145a. "Research cannot flourish if data are not preserved and made accessible. Data management should be woven into every course in science." - Nature 461, 145

Figura 27. Imagen original via doi: 10.1038/461145a. La investigación no puede florecer sin preservar los datos y hacerlos accesibles. La gestión de datos es necesaria en cada entramado de la ciencia³⁹.

Tres elementos son clave en relación con los datos de biodiversidad: la disponibilidad, la accesibilidad y la confiabilidad de los datos (Turner *et al.*, 2015). La accesibilidad es el elemento que liga con el *open data*. Las nuevas políticas de datos que promueven el acceso libre o gratuito a las imágenes satelitales del Gobierno están ampliando el uso de ciertas imágenes, aunque el número de conjuntos de datos abiertos y libres sigue siendo muy limitado. El acceso a los datos aborda la capacidad de los biólogos de la conservación y de los investigadores de la biodiversidad para descubrir, recuperar, manipular y extraer valor de imágenes satelitales, así como vincularla con otros tipos de información. Las herramientas están mejorando rápidamente el acceso (Turner *et al.*, 2015).

Sin lugar a dudas, en lo que respecta a la biodiversidad y a los objetivos relacionados con la misma, la accesibilidad es clave. En el punto relacionado con la bioinformática, se analizarán los distintos instrumentos y de qué manera se ha buscado o trabajado la accesibilidad de los datos.

3.3.1.3. Open source o software con código abierto

En la misma línea de los datos abiertos y el acceso abierto, es necesario hacer una referencia a las cuestiones de *software* de código abierto. De alguna manera, es una forma de abordar las herramientas que permiten acceder y reutilizar los datos.

En efecto, ¿puede considerarse que el hecho de que los datos que se compartan, almacenen o recaben y que requieran un *software* significa que están compartidos en abierto? O mejor, ¿en qué condiciones se pueden considerar que están en abierto? Existe *software* no accesible libremente que exige adquirir una licencia que puede resultar costosa.

El *código abierto* es el *software* distribuido y desarrollado libremente. Se focaliza más en los beneficios prácticos (acceso al código fuente) que en las cuestiones éticas o de libertad que tanta relevancia tienen en el ámbito del *software* libre.

Aunque para muchos el término «libre» hace referencia al hecho de adquirir un *software* de manera gratuita, es mucho más que eso: la libertad se refiere a poder modificar la fuente del programa sin restricciones de licencia, ya que muchas empresas de *software* encierran su código, ocultándolo y restringiendo los derechos para que no acceden terceros (Código abierto, s. f.).

³⁹ Recuperado de: <https://bit.ly/2HHEZ8S>. (Consultado por última vez el 14 de marzo de 2018).

Existen, por lo tanto, dos modelos de innovación en la organización de la ciencia, como ya se analizó en el punto relacionado con los problemas que plantea la tecnología. En efecto, por una parte, existe el modelo de inversión privada basado en el sistema de propiedad intelectual e industrial junto con el modelo colectivo, en el que el mercado trabaja de forma colaborativa. Una muestra de ello se encuentra en el *open source software* (Von Hippel y Krogh, 2003). El código abierto es un *software* que se pone a disposición de todos para que se use libremente.

Existe la posibilidad de un procedimiento concreto de distribución de *software* y de datos para la comunidad investigadora. Gómez-Díaz (2015) menciona dos tipos muy diferenciados: el *software* de código abierto y el *software* libre. El código abierto no significa únicamente que se tenga acceso al código fuente, sino que la distribución del *software* de código abierto debe cumplir con estas normas y condiciones (Open Source Initiative, s. f.):

1. *Free redistribution* (libre redistribución). La licencia no debe establecer ningún tipo de restricciones a las partes a la hora de vender o entregar el *software* como parte de un componente de una distribución agregada de *software* que contiene programas de varias fuentes diferentes. La licencia no exigirá *royalties* ni ninguna otra tasa por dicha venta.
2. *Source code* (código fuente).
3. *Derived works* (trabajos derivados).
4. *Integrity of the author's source code* (integridad del código fuente del autor).
5. *No discrimination against persons or groups* (no debe discriminar a personas ni grupos).
6. *No discrimination against fields of endeavor* (no debe discriminar ámbitos de esfuerzos).
7. *Distribution of license* (distribución de la licencia).
8. *License must not be specific to a product* (la licencia no debe ser concreta para un producto).
9. *License must not restrict other software* (la licencia no debe limitar otros *softwares*).
10. *License must be technology-neutral* (la licencia debe ser tecnológicamente neutra).

El *software* libre, por su parte, exige que los usuarios tengan las cuatro libertades esenciales (Free Software Foundation, 2001):

- *Libertad 0.* La libertad de ejecutar el programa como se desee, con cualquier propósito.
- *Libertad 1.* La libertad de estudiar cómo funciona el programa y cambiarlo para que haga lo que el usuario quiera. El acceso al código fuente es una condición necesaria para ello.
- *Libertad 2.* La libertad de redistribuir copias para ayudar a terceros.
- *Libertad 3.* La libertad de distribuir copias de sus versiones modificadas a terceros. Esto permite ofrecer a toda la comunidad la oportunidad de beneficiarse de las modificaciones. El acceso al código fuente es una condición necesaria para ello.

Aunque responden a filosofías diferentes, Gómez-Díaz (2015) los relaciona dado que las licencias más usadas incluyen ambos conceptos. En términos legales, estas licencias son contratos que participan en la construcción del marco jurídico en el que se permite la utilización, la copia, la modificación y la redistribución del *software*. El *software* libre o de código abierto no está libre (o exento) de derechos: los derechos de autor no van a desaparecer porque el *software* sea libre o de código abierto; de hecho, las licencias extienden el marco jurídico inicial definido por la ley.

Gómez-Díaz (2015) establece un procedimiento de distribución de los datos de investigación dando una lista de elementos. Los ítems marcados con un asterisco deben revisarse regularmente en las versiones importantes.

- Elegir un nombre o título para identificar el objeto, evitar las marcas y otros nombres registrados; se puede asociar a dicho nombre la fecha y el número de versión.
- (*) Establecer la lista de los autores y sus afiliaciones. Asociar un porcentaje de participación, completar con los contribuidores menores. Si la lista es larga, mantener la información al día en una página web.
- (*) Establecer la lista de las funcionalidades más importantes.
- (*) Establecer la lista de los programas o datos incluidos, e indicar las licencias o todo documento que permita su acceso, copia, modificación o redistribución.
- Elegir una licencia, con el acuerdo de todos los autores o titulares de los derechos de autor y que este acuerdo esté firmado, si es posible. Considerar la utilización de licencias libres y licencias CC (v4.0) para los datos, por ejemplo.

- Elegir una página web, *forge* o depósito para distribuir el producto, con las licencias y las condiciones de uso, copia, modificación o redistribución claramente mencionadas, así como la mejor forma de citar el trabajo. Tanto unos buenos metadatos como el respeto a estándares abiertos van a ser importantes en la distribución de nuevos componentes a una comunidad numerosa: ayudar a los demás a utilizar el trabajo de los investigadores e incrementar su longevidad. Dar licencias a la documentación (GNU FDL, CC, LAL, etc.) y a las páginas web. Utilizar identificadores persistentes siempre que sea posible.
- (*) Archivar regularmente ficheros tar.gz para trazar funcionalidades nuevas e importantes.
- Informar a los laboratorios e instituciones de los autores (si no se hizo en el paso de licencia).
- Crear e indicar claramente una dirección de contacto.
- Distribuir los datos.
- Informar a la comunidad, considerar la publicación *Data papers*.

La autora, también hace referencia a las licencias FLOSS (estas licencias están relacionadas con el acceso libre) y las licencias FOSS (relacionadas con el acceso abierto) haciendo una clara diferencia entre ambas.

Las ventajas de las licencias FOSS es que son permisivas, es decir, que permiten la creación y distribución de trabajos derivados sin restricciones. Además, son muy adecuadas para los desarrolladores de *software*. Por último, no plantean problemas de compatibilidad de la licencia, con independencia de que se hable de licencias *copyleft* o de licencias propietarias (Foss Lawyers Institute, s. f.).

Sin embargo, no son todo ventajas: las obras derivadas pueden convertirse en propietarias y afectar, con ello, a la libertad futura de los usuarios —en términos de la Free Software Foundation— (Foss Lawyers Institute, s. f.).

En lo que respecta a la ventaja que aporta la licencia FLOSS, es que reduce los costes totales del producto y elimina, también, la dependencia de los vendedores de *software*. Por lo tanto, proporciona seguridad y estabilidad (Montes, 2011).

Después de estas reflexiones previas de las cuestiones que se plantean, se va a proceder a analizar el impacto que ello ha tenido en dos entornos diferenciados: por una parte, en Europa y, por otra, en Estados Unidos.

3.3.2. El modelo europeo

Existe una tendencia innegable en la Unión Europea de desarrollo del OA. Dicha tendencia se ha ido consolidando en los diferentes programas de financiación de la Unión Europea, como puede ser el 7.º Programa Marco, en proyectos concretos y de forma no muy ordenada, es decir, que se ha ido reflejando en distintos documentos y programas.

Los programas marco son la principal iniciativa comunitaria de fomento y apoyo a la I+D dentro de la Unión Europea. Su objetivo principal es la mejora de la competitividad mediante la financiación, fundamentalmente, de actividades de investigación, desarrollo tecnológico, demostración e innovación en régimen de colaboración transnacional entre empresas e instituciones de investigación pertenecientes tanto a los países de la Unión Europea y Estados asociados como a terceros países.

Además, presta apoyo financiero a la mejora y coordinación de las infraestructuras de investigación europeas, a la promoción y formación del personal investigador, la investigación básica y, especialmente, a partir del actual 7.º Programa Marco, a la coordinación de los programas nacionales de I+D y a la puesta en funcionamiento de plataformas tecnológicas europeas concebidas para promover agendas estratégicas de investigación en sectores clave con el concurso de todos los actores implicados. Además de las plataformas tecnológicas europeas, a nivel nacional, se están promoviendo las plataformas nacionales (Centro para el Desarrollo Tecnológico Industrial, s. f.).

El 7.º Programa Marco (Comisión Europea, s. f.) señalaba en su introducción que toda investigación arranca de trabajos previos y que, por lo tanto, dependía de la capacidad que tuvieran los científicos para acceder o compartir la información científica. Es evidente que la llegada de Internet y, con ella, la posibilidad de realizar publicaciones electrónicas han abierto un sinnúmero de posibilidades de promover y mejorar la diseminación del conocimiento, con lo que se logra una mayor eficiencia de los descubrimientos científicos y una maximización del retorno de la inversión de la I+D financiada con fondos públicos.

La Unión Europea quiso usar el 7.º Programa Marco como proyecto piloto de acceso abierto a los artículos revisados por pares. Dentro del piloto, se exigía dar acceso a los artículos que tuvieran su origen en investigación financiada durante un determinado periodo de tiempo. El piloto finalizó con la finalización del 7.º Programa Marco.

Este tuvo impacto en las áreas de medioambiente, energía, salud, tecnologías de la información y del conocimiento, e-infraestructuras, ciencia en la sociedad, ciencias socioeconómicas y humanidades. Respecto a los estudios realizados en estos

ámbitos, fue necesario incluir en el acuerdo de la subvención una cláusula especial con la finalidad de que se depositasen todos los artículos en un repositorio institucional. No obstante, la Comisión ofreció la posibilidad a todos los que publicasen trabajos dentro del 7.º Programa Marco de acogerse al *open access* y publicar los resultados en el repositorio institucional. La contrapartida fundamental obtenida fue una mayor visibilidad y diseminación de los trabajos publicados.

Dentro del paraguas del 7.º Programa Marco, proliferaron proyectos cuya finalidad era precisamente ver la forma en que se puede desarrollar el OA. Un ejemplo de ello fue el proyecto Infraestructura de Acceso Abierto para Investigación en Europa (conocido como OpenAIRE, por las siglas en inglés de Open Access Infrastructure for Research in Europe) (Schmidt y Kuchma, 2012).

OpenAIRE es un proyecto con una duración inicial prevista de tres años (2009-2012) cuya finalidad era crear una red de repositorios abiertos que facilitasen un acceso abierto gratuito al conocimiento obtenido en el marco de las investigaciones subvencionadas por la Comisión Europea o por el Consejo Europeo de Investigación (CEI).

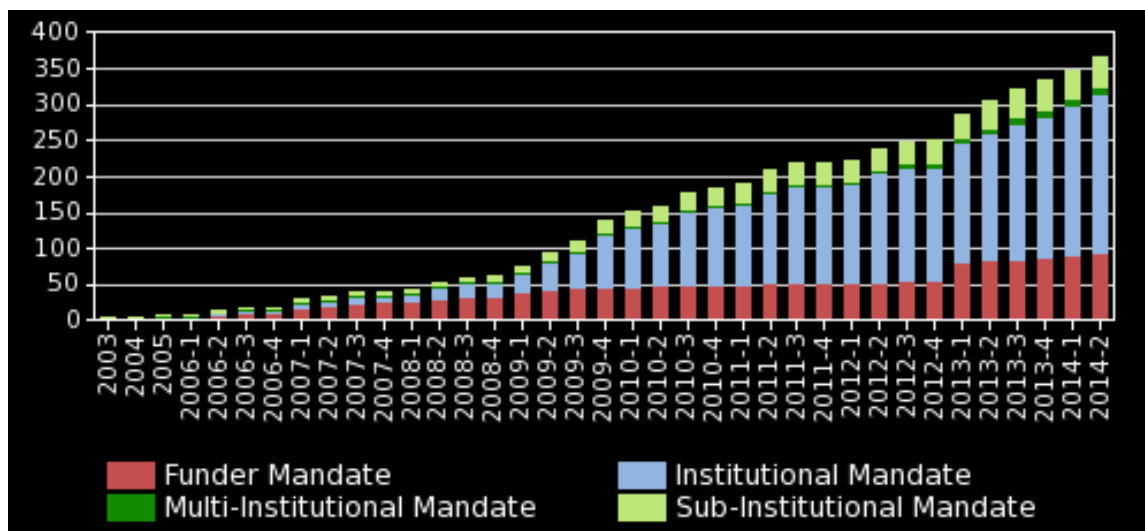


Figura 28. Crecimiento de los mandatos de acceso abierto (basado en el roadmap desde el 25 de mayo de 2012). Fuente: ROADMAP Schmidt y Kuchma, 2012.

La Comisión Europea puso de manifiesto sus intenciones, sus propósitos en relación con la ciencia y la innovación. Así, uno de sus objetivos era la optimización del impacto de la investigación financiada con fondos públicos tanto a nivel europeo (como puede ser en los programas marco 7º Programa Marco o, ahora, Horizonte 2020) como a nivel nacional de cada uno de los Estados. Era fundamental para que Europa pudiera potenciar su desempeño económico y mejorar la capacidad de competir con el conocimiento. Considera el OA una de las maneras de llegar a ello.

De esta manera, puede obtenerse una mayor y más rápida diseminación de los resultados de la investigación financiada con fondos públicos, lo que supone un beneficio para los investigadores, la industria innovadora y los ciudadanos. El OA puede estimular, también, la visibilidad de la investigación europea y, en particular, ofrecer a las pequeñas y medianas empresas acceder a lo último en investigación.

Las exigencias de *open access* encuentran su fundamento en un apoyo equilibrado tanto al OA verde (*open access* inmediato o retardado facilitado a través del autoarchivo) y OA dorado (*open access* inmediato facilitado por una editorial).

La estrategia de la Comisión es, también, impulsar las iniciativas nacionales a nivel de Estado miembro y contribuir a su coordinación dentro del área europea de investigación. Durante el Horizonte 2020, la Comisión sigue comprometida con los *stakeholders* promoviendo, al mismo tiempo, una cultura de compartir las publicaciones científicas y con el respeto debido a los derechos de los afectados, los datos de investigación o datos científicos.

En este marco se publicó la Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones (Comisión Europea, 2011) que hizo hincapié en la importancia de los recursos y, en particular, de los datos públicos, es decir, toda la información que generasen, recogiesen o sufragasen los organismos públicos de la Unión Europea. Este informe pone de relieve la fuerza que está cobrando el tema en varios Estados de la Unión Europea; sin embargo, el desarrollo no está siendo uniforme.

Un informe posterior de la Unión Europea (Comisión Europea, 2013) comenzó a diseñar las características de la futura economía de los datos y estableció algunas conclusiones operativas para apoyar y acelerar la transición. Asumió, en primer lugar, la lentitud con la que la Unión Europea está adaptándose a la actual revolución de los datos. Reconoció también, que, en la actualidad, no tenía la capacidad para responder de la misma manera y que era necesario abordar esto desde la confianza. La economía de los datos exige un elevado nivel de confianza. Por supuesto, el hecho de construir formas para que los datos circulen libremente no es óbice para que se mantengan los niveles de protección de los mismos en los casos en los que sea necesario, por ejemplo: en lo que respecta a la protección de datos de carácter personal.

El informe definía los datos abiertos como el subconjunto de datos puestos libremente a disposición de cualquiera para su reutilización tanto con fines comerciales como no comerciales (Comisión Europea, 2013). Señalaba, asimismo, la necesidad de construir una economía de los datos. Para que existiese, deberían concurrir una serie de condiciones:

- La disponibilidad de conjuntos de datos de buena calidad, fiables e interoperables y de e-infraestructuras que lo posibiliten.
- Mejores condiciones marco que faciliten la generación de valor a partir de los conjuntos de datos.
- Diversas áreas de aplicación en las que un mejor procesamiento de los macrodatos puede marcar la diferencia.

Daba un paso más allá y declaraba que era importante construir una cooperación estratégica a través de una asociación entre lo público y lo privado, de carácter contractual APPC para el desarrollo de una comunidad de datos y el fomento del intercambio de mejores prácticas. Sin lugar a dudas, además, deberán fomentarse las condiciones marco: disponibilidad de datos e interoperabilidad de los mismos.

En las conclusiones del proyecto OpenAIRE, desde que comenzara el piloto de la Comisión Europea de *open access* en el mes de agosto de 2008, se han incrementado notoriamente los mandatos de OA en el mundo (Schmidt y Kuchma, 2012). OpenAIRE cifra este incremento en un 400 %. En total, indica, el número de mandatos de OA casi se ha duplicado entre los investigadores subvencionados y se ha multiplicado en más de nueve veces entre las instituciones. Ha servido para que algunas universidades e instituciones se replanteen sus políticas de acceso abierto.

En agosto de 2016 tuvo lugar una modificación en relación con los datos abiertos en Horizonte 2020 que exigía que todos los datos derivados de la investigación realizada con estas subvenciones se pusieran en abierto. De alguna manera, representa un acercamiento al modelo estadounidense de políticas generales de *public access*.

3.3.3. El marco legislativo y las iniciativas políticas

Existen una serie de iniciativas que afectan a la construcción de este marco legislativo en las que subyace la preocupación de ir potenciando la tendencia de poner datos en abierto en la Agenda Digital de Europa. En efecto, este documento, de 2010, la Comisión Europea enfatizó la importancia de la disponibilidad de los datos del sector público para estimular los mercados de contenido digital (Janssen, 2011).

3.3.3.1. La directiva sobre la reutilización de la información del sector público

Esta directiva que, en adelante, denominaremos Directiva PSI, establece un marco general a nivel europeo con un grado mínimo de armonización (Parlamento Europeo y Consejo de la Unión Europea, 2003) y tiene un impacto en los datos

abiertos gubernamentales (Janssen, 2011), pues los organismos públicos son los mayores generadores y recabadores de datos en muchos de los distintos ámbitos: datos geográficos, información turística, etc. Estos datos, críticos para el funcionamiento de los servicios públicos, pueden ser también muy relevantes para los ciudadanos, para las organizaciones e incluso para la industria. Por lo tanto, abrir estos datos a los distintos colectivos es clave.

Esta preocupación por abrir los datos no es nueva. La Directiva PSI ha tenido de modo incuestionable un efecto positivo en la disponibilidad de los datos del sector público, aunque su relevancia, su impacto no ha sido el mismo en todas partes. Unos países la han interiorizado mejor que otros, pues la variabilidad está muy relacionada con la renuencia de cada Estado a compartir datos.

En la exposición de motivos, la directiva señaló que la información del sector público constituyó una materia prima importante para diversos productos y servicios de contenidos digitales con un potencial enorme que, hasta ahora, no ha sido explotado.

El objetivo general de esta actuación de la Unión Europea es contribuir al crecimiento económico y a la creación de empleo, liberando el potencial económico de los datos, propiedad de las Administraciones Públicas, gracias a una mejora de las condiciones de explotación. El objetivo general se adecuaba plenamente a las estrategias horizontales de la Unión Europea y, en concreto, a la estrategia Europa 2020 de la Comisión, iniciada el 3 de marzo de 2010 con el objetivo de convertir a Europa «en una economía inteligente, sostenible e integradora que genere altos niveles de empleo, productividad y cohesión social» (Parlamento Europeo y Consejo de la Unión Europea, 2003). De hecho, una de las preocupaciones de base a la hora de desarrollar la directiva fue la falta de competitividad con Estados Unidos en el mercado de los datos.

Aunque la directiva no ha alcanzado con plenitud su objetivo de apertura de datos, ha contribuido activamente a la mejora del acceso a los datos gubernamentales (Janssen, 2011). El objetivo de la directiva quedó recogido en el artículo 1 y no era otro que establecer un conjunto mínimo de normas que regulen la reutilización y los instrumentos prácticos que facilitasen la reutilización de los documentos existentes conservados por organismos del sector público de los Estados miembros. Respondiendo a su condición de directiva, establecían unos mínimos que tenían que cumplir los Estados miembros y que, al mismo tiempo, invitasen a estos a ser proactivos en la adopción de medidas que lo facilitasen.

La directiva era de aplicación a la reutilización de los documentos existentes en poder de organismos del sector público y, por consiguiente, no obligaban a los Estados miembros a crear datos del sector público que estimularan el mercado, sino

tan solo a poner a disposición los datos que se creasen mientras desempeñasen sus tareas habituales. Afectaba a cualquier tipo de documento. La directiva definía documento en el artículo 2.3: «documento: a) cualquier contenido sea cual sea el soporte (escrito en papel o almacenado en forma electrónica o como grabación sonora, visual o audiovisual); b) cualquier parte de tal contenido» (Parlamento Europeo y Consejo de la Unión Europea, 2003).

La directiva definía el concepto de reutilización en el apartado 4 del mismo artículo 2:

Reutilización: el uso de documentos que obran en poder de organismos del sector público por personas físicas o jurídicas con fines comerciales o no comerciales distintos del propósito inicial que tenían esos documentos en la misión de servicio público para la que se produjeron. El intercambio de documentos entre organismos del sector público en el marco de sus actividades de servicio público no se considerará reutilización (Parlamento Europeo y Consejo de la Unión Europea, 2003).

El artículo 1.2 de la directiva (Parlamento Europeo y Consejo de la Unión Europea, 2003) recogía algunas exclusiones [el destacado es nuestro]:

2. La presente Directiva no se aplicará a:

- a. los documentos cuyo suministro sea una *actividad* que se *salga* del *ámbito* de la misión de servicio público de los organismos del sector *público* afectados, definida con arreglo a la legislación o a otras normas de obligado cumplimiento del Estado miembro o, en su ausencia, definida en consonancia con la práctica administrativa común del Estado miembro de que se trate;
- b. los documentos sobre los que existan *derechos de propiedad intelectual* por parte de terceros;
- c. los documentos a los que no pueda accederse en virtud de regímenes de acceso de los Estados miembros, por motivos, entre otros, de: — *protección de la seguridad nacional* (esto es, seguridad del Estado), *defensa* o *seguridad pública* — confidencialidad estadística o comercial;
- d. los *documentos* conservados por las entidades de radiodifusión de servicio público y sus filiales, y por otras entidades o sus filiales para el *cumplimiento de una misión de radiodifusión* de servicio público;
- e. los *documentos* conservados por *instituciones educativas y de investigación*, tales como centros escolares, universidades, archivos, bibliotecas y centros de investigación, con inclusión, si procede, de organizaciones creadas para la transferencia de los resultados de la investigación;

- f. los *documentos* conservados por *instituciones culturales* tales como museos, bibliotecas, archivos, orquestas, óperas, ballets y teatros.

Además de estas restricciones de la propia Directiva PSI, es necesario tener en cuenta las restricciones derivadas de las normas de protección de datos, en la actualidad recogidas en el Reglamento de Protección de datos (Parlamento Europeo y Consejo de la Unión Europea, 2016) recientemente publicado y, con anterioridad, en la Directiva de Protección de Datos (Parlamento Europeo y Consejo de la Unión Europea, 1995).

El artículo 3 de la directiva (Parlamento Europeo y Consejo de la Unión Europea, 2003) establece el principio general [el destacado es nuestro]:

Los Estados miembros velarán por que, cuando se autorice la reutilización de documentos conservados por organismos del sector público, dichos documentos puedan ser *reutilizados para fines comerciales o no comerciales* de conformidad con las condiciones expuestas en los capítulos III y IV. *En la medida de lo posible, los documentos se pondrán a disposición del público por medios electrónicos.*

La Directiva PSI incluye disposiciones que ayudan a los reutilizadores a encontrar los datos que necesitan y a utilizarlos sin restricciones innecesarias (Janssen, 2011).

Otra de las barreras que existe en Europa en lo que respecta al desarrollo del mercado de la información es la falta de transparencia en relación con la disponibilidad de los documentos y las condiciones con sujeción a las cuales pueden reutilizarse. A efectos de resolver tal problema, la directiva recoge en su artículo 7 las condiciones de transparencia y, en el artículo 8, que las licencias estándar están disponibles en formatos digitales en los Estados miembros y que se fomenta que los organismos públicos utilicen estas licencias estándar. Además, en el artículo 9 prevé que los Estados miembros deberán asegurarse de que existen dispositivos prácticos que faciliten la búsqueda de los documentos disponibles (Parlamento Europeo y Consejo de la Unión Europea, 2003).

Janssen (2011) señala que puede que la directiva genere cierta confusión entre los conceptos de acceso y reutilización. Esta directiva se revisa en el año 2013 (Parlamento Europeo y Consejo de la Unión Europea, 2013) y recoge la visión formulada en 2011 por la Comisión Europea. La revisión mantiene la visión de la reutilización de la información del sector público como un valor añadido. Introduce, además, una autorización generalizada a la reutilización de la información. Desde la revisión, las normas serán de aplicación a archivos, bibliotecas y museos.

Una de las novedades principales que introduce es la recomendación de publicar la información preferentemente en formatos estándares, abiertos y procesables de modo automatizado, con el mayor nivel de granularidad posible. Teniendo en

cuenta el derecho de reutilización, se recomienda asimismo no introducir restricciones o, si se introducen, que sean el menor número posible y que se recojan, siempre que sea posible, en una licencia que pueda usarse por medios electrónicos.

Aunque la directiva no prohíbe que se cobren tasas, sí busca que se limiten lo máximo posible. Podrán seguir firmándose acuerdos exclusivos en materia de reutilización de la información por razón de interés público, por lo general, limitándose a tres años.

Además, existen otras iniciativas políticas en ámbitos concretos como la Directiva 2003/4/CE (Parlamento Europeo y Consejo de la Unión Europea, 2003) y otras que se analizarán a lo largo del presente trabajo.

3.3.3.2. La Directiva INSPIRE

La finalidad de esta directiva firmada por el Parlamento Europeo y el Consejo de la Unión Europea (2007) —que en adelante denominaremos Directiva INSPIRE— fue construir una infraestructura de datos espaciales con la finalidad de implementar políticas medioambientales así como actividades que puedan tener un impacto medioambiental. La idea es permitir el intercambio de información espacial medioambiental dentro de las organizaciones del sector público y facilitar el acceso público a la información espacial dentro de Europa contribuyendo a la creación de políticas y normas más allá de las fronteras. En relación con el medioambiente no existen fronteras, pero es importante encontrar estándares de registro y formas de compartir los datos para poder tomar decisiones informadas en el momento preciso. Un ejemplo de ello fue la gestión de la erupción del volcán en Islandia. Los gases que se desprendieron afectaron a otros países más allá de las del país fronteras, lo que hizo necesario evaluar los niveles existentes de contaminación y el impacto que ello pudiera tener.

Una de las exigencias de la Directiva INSPIRE es la creación de un Geo Portal para compartir los datos y eliminar todos los obstáculos que sean posibles (considerandos 19 y 20). Por lo tanto, la Directiva INSPIRE exige a las autoridades públicas en toda la Unión Europea que den acceso a sus conjuntos de datos medioambientales de manera uniforme e interoperable. Para garantizar este nivel de interoperabilidad, la directiva exigió que se adoptaran normas de implementación comunes para los metadatos, las especificaciones de datos, los servicios de red y el intercambio de datos dentro de la infraestructura paneuropea de datos espaciales (Shaon y Woolf, 2011).

Por lo general, señalan Shaon y Woolf (2011), el acceso sostenible a los datos medioambientales es cada vez más importante y, al mismo tiempo, cada vez más difícil. La información espacial a escala local, regional y global es la base de la

geografía. De hecho, fruto del cambio que se ha producido en la tecnología, la atención se pone ahora en los datos globales disponibles para todos, de forma increíblemente abundante. Solo las formas actuales de gestionar la información a través de la tecnología y la tarea de compartir pueden ayudar a gestionar esta gran cantidad de datos (Bartha y Kocsis, 2011). Los cinco elementos que integran INSPIRE (Bartha y Kocsis, 2011) son los que se reflejan en el siguiente esquema.

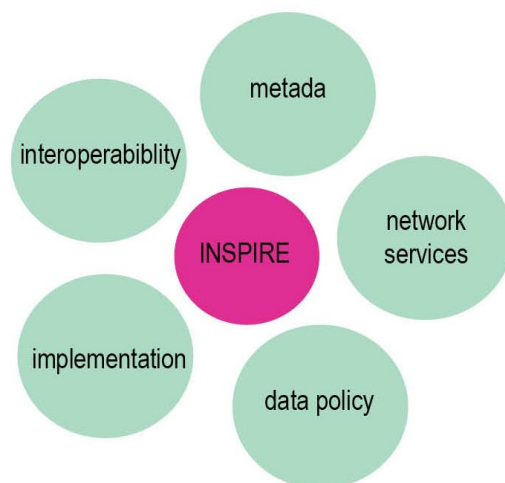


Figura 29. Los cinco elementos de INSPIRE. Fuente: Bartha y Kocsis, 2011.

Como algún autor puso de manifiesto, la Directiva INSPIRE es innovadora en dos aspectos: el primero, porque trató de construir una infraestructura basándose en lo que ya existía en los distintos ámbitos (nacional, regional, etc.); el segundo, porque intentó comprometer de una forma mucho más estructurada a las comunidades usuarias y a los *stakeholders* de información geográfica, organizándolos a través de comunidades interesadas en los datos espaciales (Craglia y Annoni, 2006).

Es una directiva cuyo foco se encuentra en los datos geográficos, dado que la falta de interoperabilidad de los mismos dificulta la superposición de fronteras y territorios. Ejemplo que ilustra esto es lo que ocurre en España, donde existen más de ocho mil municipios con sistemas no interoperables. Esta falta de interoperabilidad genera que al acercarse a los lugares colindantes no exista claridad en lo que respecta a las delimitaciones, carreteras, caminos... Similar confusión produce la existencia de una multiplicidad de líneas costeras no interoperables que dificultan, por lo tanto, la gestión de las costas e incluso de las fronteras.

A fin de cuentas, la infraestructura de datos espaciales es un marco de políticas, estándares y tecnología que permite a los proveedores de datos publicar y a los usuarios acceder, integrar y distribuir información geoespacial heterogénea (Hansen *et al.*, 2011).

La Unión Europea ha dado un paso más con la publicación de las directrices de aplicación al Programa Horizonte 2020 (sucesor del 7.º Programa Marco). Las directrices de *open access* a las publicaciones científicas y a los datos procedentes de la investigación en Horizonte 2020 (Comisión Europea, 2016) definen el *open access* como la práctica de proporcionar acceso en línea a la información científica sin coste para el usuario final y de manera que este pueda reutilizarla.

Dentro del contexto de la investigación y la innovación, la «información científica» incluye tanto (1) los artículos de investigación revisados por pares como (2) los datos de investigación —los datos que están detrás de las publicaciones, los datos curados o los datos primarios—. Las directrices recogen asimismo la diferenciación del *open access* verde y *open access* dorado que, como ya se ha visto, son dos caminos de *open access*.

Es evidente que las tendencias de *open access* son imparables. La problemática está en cómo se van impulsando y la velocidad de las mismas. También es clave tener en cuenta cómo las distintas políticas y tendencias locales, regionales y mundiales de *open access* van encajando.

En el marco de los programas de investigación de la Unión Europea, en particular los incluidos dentro del Horizonte 2020 desde agosto de 2016, se impuso como exigencia que todos los datos fuesen publicados en abierto, con la finalidad de garantizar su reutilización y que no existieran barreras para su uso. Ya en 2012 la Comisión Europea publicó la política de acceso abierto relativa a Horizonte 2020.

En enero de 2017 volvió a actualizarse la política de Horizonte 2020. Partieron de la formulación del reto de base: un acceso más amplio a los datos científicos para ayudar a los investigadores, innovadores y al público en general a encontrar y reutilizar los datos y a verificar los resultados.

A fin de cuentas, esto supone un mejor retorno de la inversión de los fondos con los que la Unión Europea subvenciona la investigación, generando, de esta manera, un mayor beneficio público. Además, sirve para promover la investigación entre las distintas disciplinas científicas, hoy clave, para resolver los complejos cambios sociales a los que hay que hacer frente (Comisión Europea, 2017).



Figura 30. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.

Como consecuencia de todo lo anterior, la conclusión es que Horizonte 2020 está recomendando u ordenando que todas las publicaciones científicas se hagan en *open access* con la finalidad de que puedan ser reutilizadas. A partir de enero de 2017, todas las publicaciones de Horizonte 2020 serán *open access* por defecto, sin perjuicio de que el titular de derechos tenga la posibilidad de establecer que no se aplique la norma general de acceso abierto (Comisión Europea, 2017).

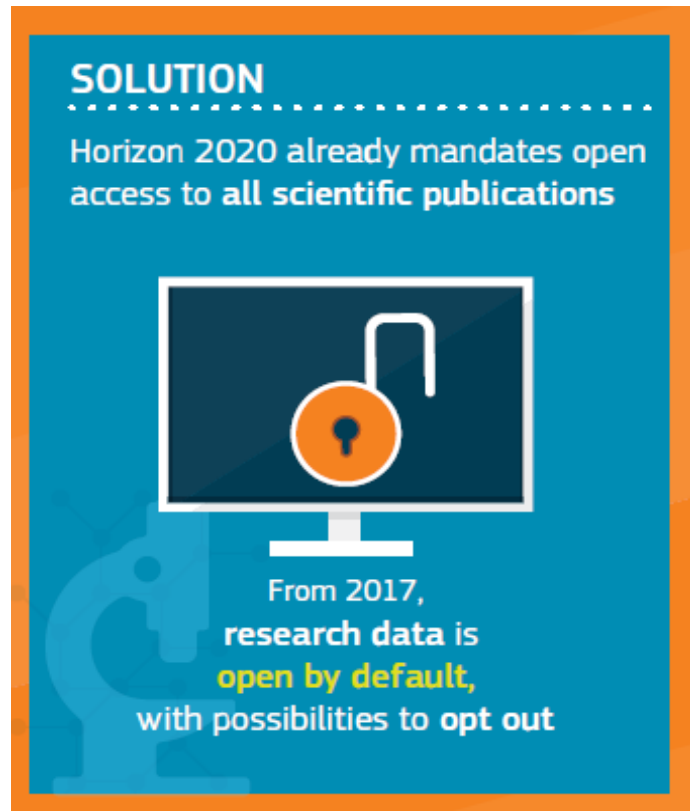


Figura 31.
2017.

Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea,

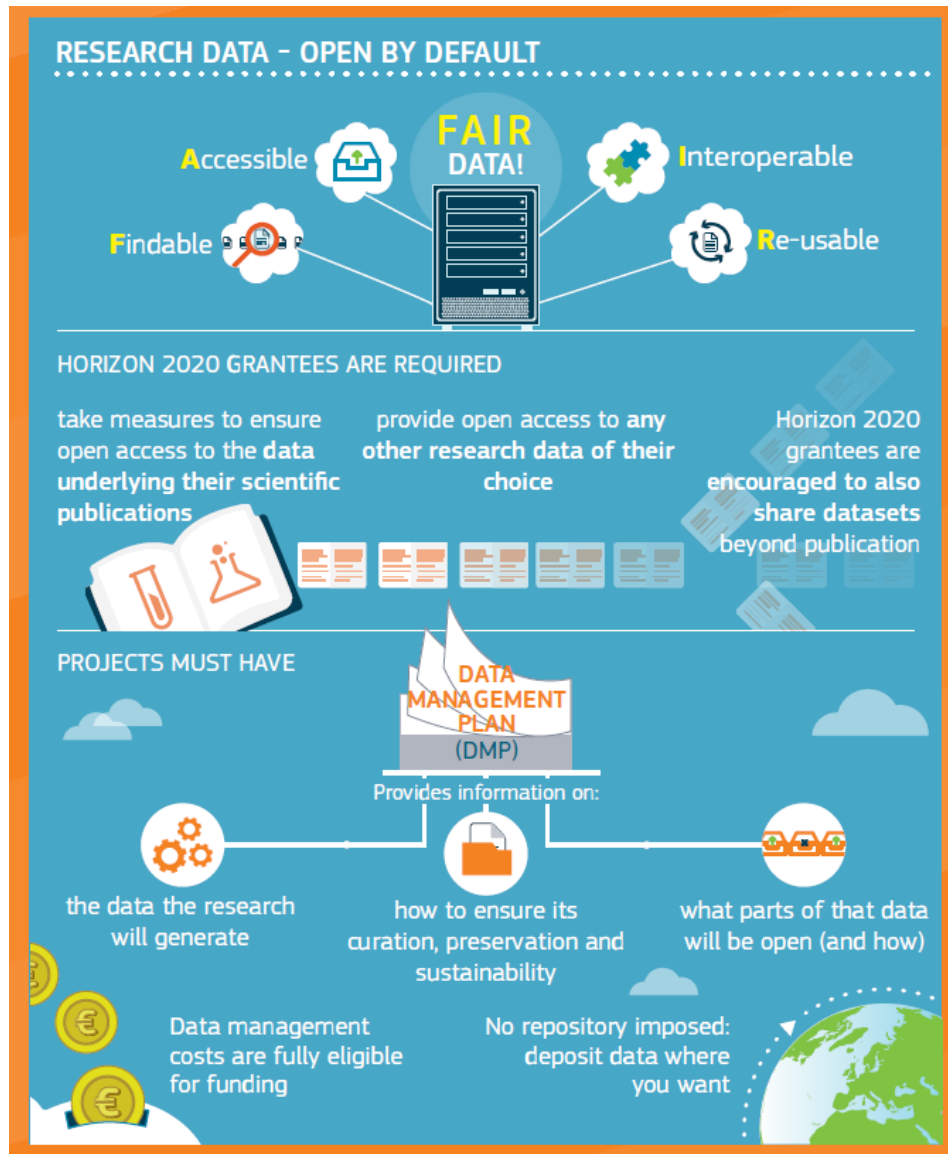


Figura 32. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.

La Comisión hace toda una reflexión de lo que supone el acceso abierto por defecto. En efecto, si bien se recoge la posibilidad de que no se aplique *open access* a los datos, los que hayan recibido la financiación tendrán que justificar la razón por la cual consideran apropiado no hacerlo.

Se recogen tres razones principales en las que se pueden amparar para excluir el acceso abierto. La primera de ellas es la privacidad; la segunda, la existencia de derechos de propiedad intelectual y, la tercera, más genérica, que pueda afectar al objetivo principal del proyecto. Por lo tanto, deberá buscarse un equilibrio entre el acceso abierto y la protección total.

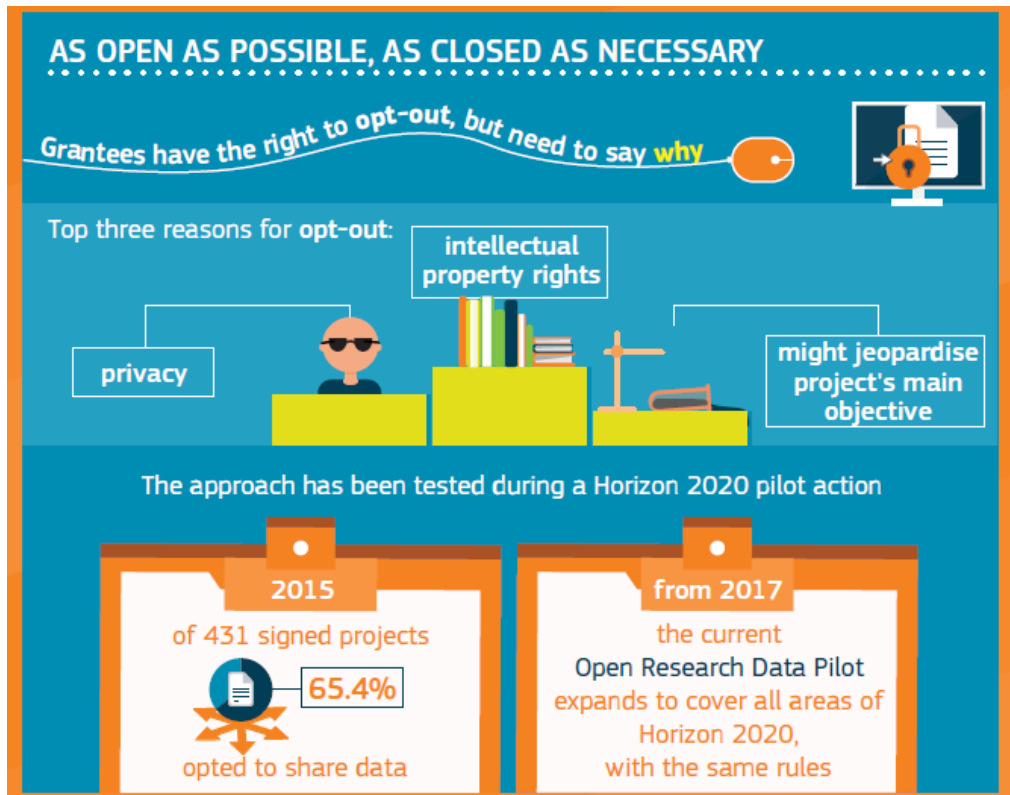


Figura 33. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.

Además, la Comisión analiza de qué manera funciona. Lo importante, a fin de cuentas, es que se publiquen los resultados, elegir un repositorio donde se publiquen los datos e informar a OpenAIRE, el portal de acceso abierto de la Unión Europea financiado con fondos europeos.

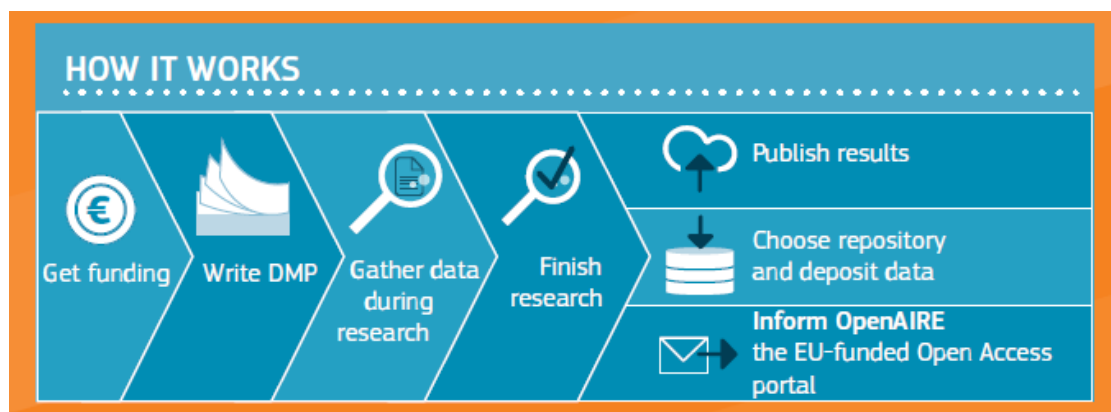


Figura 34. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.

Todo este proceso, es parte del movimiento en el que la ciencia se encuentra inmersa y que consiste en compartir datos, abrirlos con la finalidad de contribuir a la aceleración de la investigación.

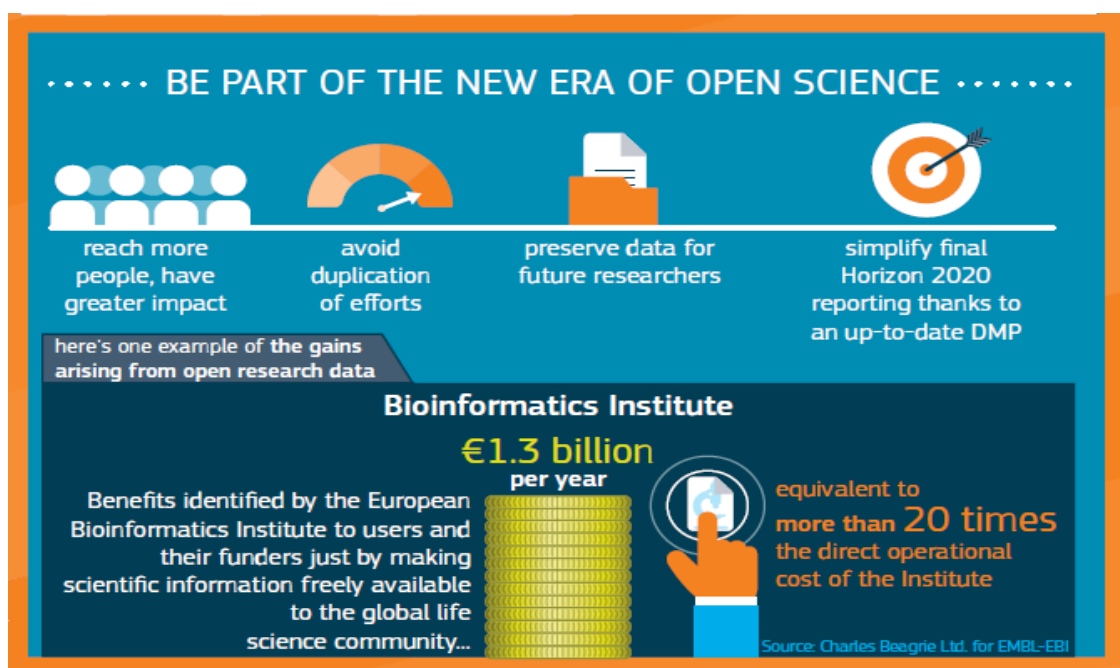


Figura 35. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.

La finalidad de la ciencia abierta es, sin lugar a dudas, llegar a más personas y tener más impacto. Además, persigue no duplicar esfuerzos y preservar los datos para los futuros investigadores, así como simplificar el reporte de los proyectos financiados con fondos europeos. Existe una multiplicidad de estudios sobre los beneficios derivados de la *open science*.

En el artículo 37 de la Ley de la Ciencia española (Anon. 2011) se recoge la idea de las publicaciones de los datos en abierto: todo lo financiado con fondos públicos españoles y europeos estarán abiertos salvo que el científico lo restrinja para su publicación.

3.3.3.3. El mercado único digital

En el año 2014 arrancó una reforma del derecho de propiedad intelectual europeo con ambiciosos objetivos de modernizar las normas de propiedad intelectual de la Unión Europea con la finalidad de crear un mercado único digital.

Sin embargo, hoy se ha convertido en un esfuerzo fragmentado de proteger los autoproclamados derechos de los titulares de derechos (Keller, 2016). La Comisión, anteponiendo la protección de los derechos de los titulares a la armonización las normas de derecho de propiedad intelectual en Europa y asegurándose de que no limitarán el potencial de las nuevas tecnologías, ha optado por centrarse en apoyar los modelos empresariales heredados de los titulares de derechos en un intento por defender el *status quo*. Este es un enfoque notablemente miope para modernizar uno

de los marcos políticos básicos que gobiernan la economía de la información, especialmente si se tiene en cuenta que estas reglas probablemente seguirán vigentes durante las próximas décadas (Keller, 2016).

Parece que la Comisión considera que las licencias son la nueva solución. Aunque no pretende establecer la necesidad de licencias para todo, será necesario licenciar, por lo tanto, pedir autorización para muchos de los usos.

Parece poco consistente con todas las reflexiones se hayan ido realizando hasta ahora en relación con la eliminación de las barreras. De hecho, como ya se ha analizado hasta ahora y se verá, también, después, dentro de la Unión Europea existen al mismo tiempo toda una serie de iniciativas que persiguen el acceso abierto.

La Comisión ha presentado un borrador de directiva de propiedad intelectual para la creación del mercado digital único. La idea que tenía era que dicha regulación se aprobase en el año 2017.

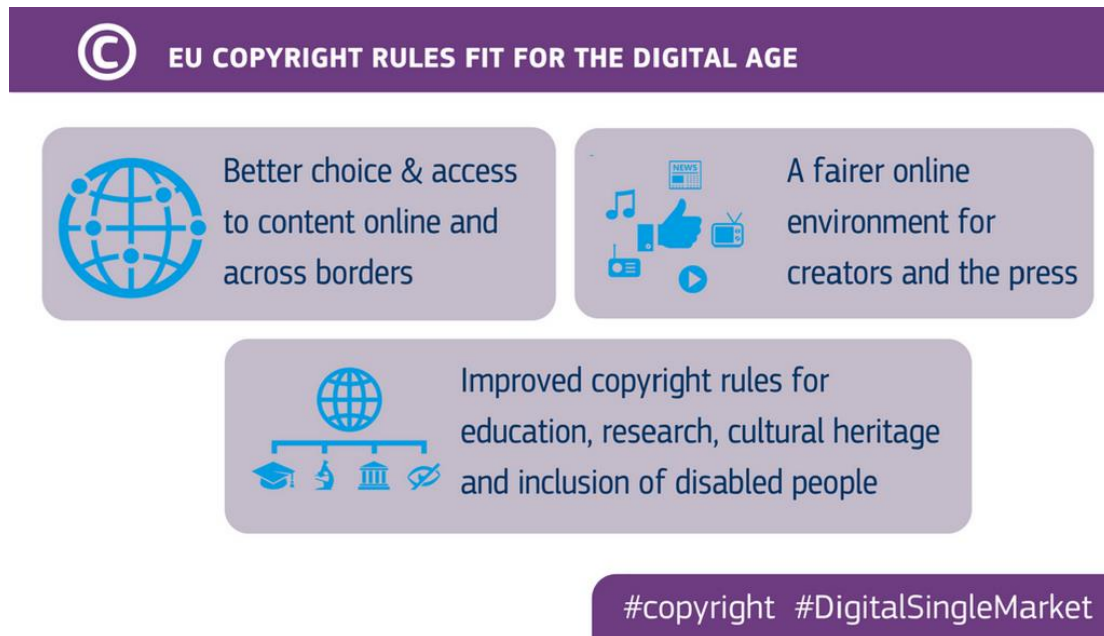


Figura 36. Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2016.

Esta diapositiva publicada por la Comisión Europea (2016) recoge los principios sobre los que quiere articular esta reforma. En primer lugar, es clave que exista tanto un mejor acceso como una mejor elección de los contenidos en línea, que no tenga en cuenta las fronteras, en relación con la libre circulación dentro de la Unión Europea.

En segundo lugar, es importante la creación de un entorno digital más equitativo tanto para los creadores como para la prensa. Por último, se deberían mejorar las

normas de derecho intelectual en relación con los sectores de la educación y de la investigación, así como con el patrimonio cultural y con los discapacitados.

3.3.3.4. La Agenda Digital para Europa

La Unión Europea ha elaborado una Agenda Digital para Europa, una de las iniciativas clave de la Estrategia 2020 para lograr un crecimiento sostenible, inclusivo y eficiente (Comisión Europea, 2010). Estos son los obstáculos que dificultan dicha Agenda Digital:

- La fragmentación de los mercados digitales.
- La falta de interoperabilidad.
- El incremento de la ciberdelincuencia y el riesgo de escasa confianza en las redes.
- La ausencia de inversión en las redes.
- La insuficiencia de los esfuerzos de investigación e innovación.
- Las carencias en la alfabetización y la capacitación digitales.
- La pérdida de oportunidades para afrontar los retos sociales.

La economía europea exige centrarse en el conocimiento y en la innovación (Mansell, 2014) y, al mismo tiempo, la economía digital está creciendo a un ritmo siete veces superior al crecimiento de la economía normal (Comisión Europea, 2014). Así, la finalidad de la Agenda Digital era poner remedio a los obstáculos existentes y, para ello, quería trabajar a distintos niveles.

Una de las claves que tiene en cuenta dicha agenda es el ancho de banda, pues la velocidad en las comunicaciones es hoy igual de clave que, en su momento, lo fueron el desarrollo del transporte o de la electricidad.

Otra de las claves es, sin duda, la creación de un mercado único europeo de telecomunicaciones. Para la construcción de este mercado único se están promoviendo toda una serie de medidas como puede ser la supresión del *roaming* en todo el territorio de la Unión Europea o la coordinación de los espectros de comunicación.

También es fundamental contribuir al desarrollo de las *startups*. Este exige un acceso más accesible pero, también, una mejor protección contra los ciberataques. Todo esto lo resume la comisaria N. Kroes (s. f.) en el siguiente gráfico.

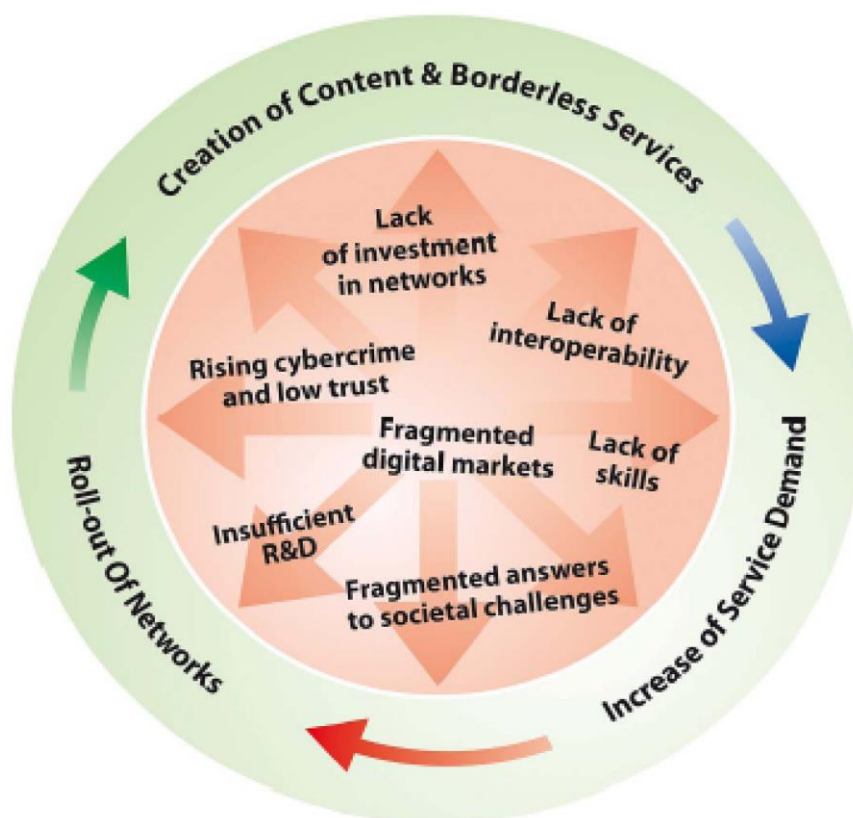


Figura 37. Fuente: N. Kroes.

Como ya se ha puesto de manifiesto, tanto el acceso abierto como la interoperabilidad son factores esenciales para la construcción de la Agenda Digital europea.

3.3.3.5. La licencia sobre el *text and data mining*: ¿un paso atrás?

Análisis del marco legislativo y las iniciativas políticas de *open access* y sus contradicciones internas con las licencias TDM.

La Comisión reconoce el increíble potencial que tiene la minería de texto y de datos. En efecto, la propia Comisión señala que la minería de texto y datos puede ser una herramienta científica de investigación muy poderosa para analizar grandes cantidades de texto y de datos como pueden ser los supuestos de publicaciones científicas y de conjuntos de datos de investigación.

De la misma manera, pone de relieve que sería más fácil que los investigadores recurrieran a este tipo de herramientas, aunque introduce un matiz: la existencia de incertidumbre legal en relación con los derechos de autor.

El borrador de la directiva pone de manifiesto que, hoy, existen disposiciones del derecho europeo que ya cubren algunas de las actividades de minería de texto y datos, con la salvedad de que estas excepciones son opcionales y no están

totalmente adaptadas al uso que la investigación científica está haciendo, actualmente, de la tecnología.

Con la finalidad de superar esta incertidumbre jurídica, el borrador de directiva introduce una excepción obligatoria en lo que respecta a los usos de las tecnologías de minería de texto y datos en el ámbito de la investigación científica (Vollmer, 2016).

El artículo 3 del borrador de la directiva establece que los Estados miembros deberán incorporar una excepción a los derechos exclusivos que establece la directiva de propiedad intelectual en lo que respecta a las reproducciones y extracciones que realicen las organizaciones de investigación, con la finalidad de realizar minería de texto y datos de trabajos u otros temas a los que tengan acceso, de conformidad con la ley y a los efectos de llevar a cabo una investigación científica. Cualquier disposición en contrario será nula (Vollmer, 2016).

Se plantean algunas cuestiones en relación con este enfoque. Al hacer la directiva obligatoria se aseguran de que la excepción aplique de manera uniforme en todos los Estados miembros (Vollmer, 2016). Además, es importante, la aclaración explícita de que los derechos que garantiza la excepción no pueden cederse de ninguna otra manera.

Sin lugar a dudas, es acertado que la excepción propuesta no vaya a ser de aplicación a los trabajos relacionados con la Administración. Sobre todo, habida cuenta de que, con anterioridad, la Comisión solo incluía, en la excepción, usos de investigación sin fines comerciales. Aunque, por el contrario, no quedan corregidos los defectos de que adolece.

Introduce una excepción que solo puede favorecer a un reducido grupo de beneficiarios (en este caso: organizaciones de investigación). Queda claro que cualquier otra persona necesitará obtener el permiso de los titulares del derecho antes de proceder a realizar minería de texto y datos.

En otras palabras: lo que la Comisión está proponiendo es que cada acto de minería de texto y de datos (que suponga una reproducción o extracción) exigirá una licencia salvo que la realice una organización de investigación. La pregunta que surge es: ¿existen actos de minería de texto y de datos que no supongan reproducción ni extracción? Parece que la respuesta es claramente que no.

Sin lugar a dudas, esta medida, si realmente se aprueba, va a erigir un sinnúmero de barreras para organizaciones que no sean de investigación, como pueden ser las relacionadas con las ciencias de la vida o la farmacéutica. La excepción que elimina el uso comercial, salva a este sector, no así al resto.

Es más, esto va a significar que algunas fuentes de datos quedarán fuera de los límites de minería de texto y de datos para cualquiera que no trabaje en organizaciones de investigación. Grandes cantidades de los materiales almacenados en Internet públicamente o en archivos no están disponibles para ser licenciados, dado que la propiedad de los mismos está fragmentada o los titulares de los derechos han desaparecido o no tienen ningún incentivo para dar licencias de sus trabajos.

Existe una extendida creencia de que el origen de las restricciones y la inclusión de las licencias TDM es una reacción de la Unión Europea a las políticas de Google, Amazon y otros gigantes tecnológicos. La inclusión de este tipo de restricciones extrapoladas al mundo *offline* sería lo mismo que prohibir o no permitir al comprador de un libro volver a mirarlo para buscar referencias al mismo. Si en el contexto *offline*, no parece que esto tenga sentido, es evidente que tampoco lo tiene en la red. Abundando en esto mismo, más grave es lo que ocurre en las revistas. En efecto, la suscripción a la revista *online* supone poder acceder a la misma; sin embargo, lo que las editoriales están haciendo es restringir el acceso a las mismas si se produce cualquier modificación de formato. Es decir que la suscripción no permite el acceso a lo que he comprado en cualquier momento, sino solamente durante el periodo que ellos determinen. No parece que tenga mucho sentido este tipo de restricciones en el contexto del crecimiento ingente de los datos y de su interdependencia.

De todo esto resulta que el análisis de datos en Europa comportará unos riesgos inexistentes en otros lugares en los que la minería de texto y de datos no se considere un hecho relevante a los efectos del derecho de propiedad intelectual.

Existe una solución evidente que plantea Vollmer (2016). Es importante recordar que toda la discusión gira sobre la minería de texto y datos de materiales legalmente accesibles. Esto significa que los editores y otros titulares de derechos pueden tener en cuenta el valor de la minería de texto y dato en las licencias que otorgan a sus licenciatariaos. Esto no afectaría a sus ingresos y lo único que no permitiría sería la concesión de licencias de sus catálogos de nuevo. Visto desde este punto de vista, la propuesta de la Comisión parece concebida para permitir este tipo de doble inmersión por parte de los editores a expensas de limitar el acceso a la innovación impulsada por datos para todos los demás.

En lugar de proponer una excepción limitada, Europa necesita una excepción que no solo permita, sino que además impulse, la extracción de texto y datos de materiales legalmente accesibles por cualquier persona para cualquier propósito. Como otros han argumentado antes, puede lograrse más fácilmente mediante la inclusión de la minería de texto y datos en el ámbito de las excepciones existentes para los actos temporales de reproducción (artículo 5.1 de la Directiva InfoSoc) o

modificando la propuesta de directiva de propiedad intelectual en el mercado único digital. De manera que el artículo 3.1 se aplica a todos los usuarios y no solo a aquellos que forman parte de organizaciones de investigación.

La profesora Van Eechoud en su línea de tiempo de Twitter realiza distintas manifestaciones defendiendo la necesidad de aplicar una excepción de derechos de propiedad intelectual a los datos científicos y de investigación. Parece que la propuesta de la Comisión Europea trata de volver al momento anterior a la era digital: a los periódicos y revistas. A la luz de los cambios estructurales que se han producido en los mercados de anuncios y en los comportamientos de lectura, cabe cuestionarse si procede centrarse en la *forma* de la prensa tradicional y no en las *funciones* que cumplía.

Es más, las publicaciones en prensa, tal como quedan hoy definidas en la propuesta, cubrirían todos los ámbitos de publicaciones periódicas que sean una forma de comunicación, incluidas las publicaciones profesionales, de negocios, educativas y gubernamentales. Sin embargo, en lo que respecta a estos ámbitos, no queda muy claro que resulte necesario añadir un plus de protección de derechos de propiedad intelectual. Cuando menos, en el caso de publicaciones que tienen su origen en los órganos del sector público que deben excluirse, ya que no existe ninguna necesidad de que, en este ámbito, los derechos de propiedad intelectual incentiven las publicaciones.

De hecho, es más bien al contrario: los rápidos desarrollos hacia una mayor transparencia, la difusión activa de los sectores públicos de información y el *open licensing* (*open data*) dejan entrever que la introducción de un nuevo derecho solo contribuirá a la creación de costes y barreras adicionales. La exclusión de las publicaciones científicas y académicas tiene todo el sentido —a la vista de la fuerte posición de mercado de los editores científicos de carácter comercial y con el intento de alcanzar el *open access* y con carácter más general— con las políticas de *open science* de la Unión Europea y de sus Estados miembros. Con la finalidad de asegurar que la exclusión es inequívoca, se recomienda incluirla en una disposición sustantiva y no solo en los considerandos.

De hecho, el mundo académico y de la investigación se está posicionando claramente en la defensa de la ciencia abierta. Uno de los posicionamientos formulado en estos días por Frontier (Fenter, 2017) va, precisamente, en esta línea. Para innovar de forma efectiva, los resultados de la investigación tienen que ser abiertos para toda la sociedad. Hoy existe un crecimiento exponencial en lo que respecta al conocimiento generado en los ámbitos científico, médico y de investigación técnica, y se están desarrollando nuevas herramientas con la finalidad de que dichos datos puedan explotarse de forma poderosa. Una de las herramientas más prometedoras es la minería de datos.

Es fundamental que la Comisión Europea recoja una excepción a estos efectos que incluya a todas las organizaciones de investigación. El uso de la minería de datos debe hacerse lo más amplio y, expresamente, abierto como sea posible para beneficiar a la sociedad europea acelerando el progreso científico, la innovación y el crecimiento económico. En los próximos meses, se verá si esta regulación termina o no consolidándose y de qué manera puede afectar a la circulación de conocimiento dentro de la Unión Europea.

Es cuestionable a qué tipos de trabajos podrían aplicarse estas restricciones. En efecto, ¿cabría aplicarlo a trabajos realizados con financiación pública? ¿Tiene el titular el derecho de licenciar los trabajos? ¿Puede afectar a los datos de biodiversidad? A fin de cuentas, los datos de biodiversidad son recogidos de la naturaleza. ¿Dónde se encuentran los límites? Estas cuestiones ya se abordaron en el punto relacionado con los problemas que plantea la tecnología y el concepto de dato más arriba.

Una de las claves de la accesibilidad de los datos es la capacidad de hacer búsquedas semánticas en todo tipo de documentos para localizar datos concretos. La ciencia que permite extraer información útil de grandes conjuntos de datos o de bases de datos se conoce con el nombre de *data mining* (minería de datos). Es una nueva disciplina, en la intersección de la estadística, el aprendizaje automático, la gestión de datos y las bases de datos, los patrones de reconocimiento, la inteligencia artificial y otros muchos ámbitos (Hand *et al.*, 2001).

El crecimiento exponencial de los datos y de las bases de datos puede dificultar el acceso a la información y la capacidad de vincularla; la minería de datos sirve, por lo tanto, para observar grandes conjuntos de datos y poder extraer relaciones entre ellos.

[La minería de datos] es un proceso iterativo que se realiza de forma automática o manual, tanto más interesante cuanto que no se espera un resultado concreto (Kantardzic, 2011).

Supone una búsqueda de información nueva, relevante, no trivial y valiosa dentro de grandes conjuntos de datos. Está basada en un esfuerzo combinado entre personas y computadoras. En la práctica, sus dos objetivos primordiales tienden a ser la descripción y la predicción (Kantardzic, 2011). La predicción en cuanto que, usando una serie de variables o campos pueden predecirse, cosas desconocidas, valores futuros u otras variables valiosas (genera un modelo). Descripción en cuanto que se fija en analizar patrones que describen los datos que pueden observar las personas (aporta nueva información valiosa) (Kantardzic, 2011).

En efecto, existen importantes divergencias en el tema entre Estados Unidos y Europa que hacen difícil encontrar una postura conciliadora. Aunque pudiera

interpretarse que este tipo de protección es un freno para la compartición, no deja de ser cierto que existe, también en la Unión Europea, una tendencia cada vez más fuerte al OA. En un futuro próximo, se verá si esta licencia tiene finalmente cabida o no la tiene.

3.3.4. El modelo estadounidense: las políticas de public access

Llegados a este punto, quedan claras las cuestiones que plantean las limitaciones en el acceso a los datos. Resulta necesario analizar, sobre este punto, la evolución que se ha dado en este ámbito, en Estados Unidos, hasta llegar a las llamadas políticas de *public access*.

El movimiento de las políticas de *public access* tiene un doble origen (Suber, 2012). Por una parte, la revolución de Internet ha supuesto un crecimiento exponencial del conocimiento (Mokyr, 2002) y, por otra, un encarecimiento de las suscripciones a las revistas científicas y, en consecuencia, un encarecimiento del acceso al conocimiento. Aceptar esto, de alguna manera, supone admitir que el coste es una barrera al conocimiento, lo cual no parece tener mucho sentido ni mucha justificación. Sobre todo, teniendo en cuenta que el conocimiento tiene su origen en la investigación, por lo general, subvencionada con fondos públicos.

Los antecedentes de las políticas de *public access* pueden encontrarse en la política del National Institute of Health (NIH), primera política de *public access* en Estados Unidos. Esta política de *public access* fue un mandato de acceso abierto que exigió que las publicaciones de investigación financiadas con fondos del National Institute of Health estuviesen disponibles para el público en general en *PubMed Central* (NIH Public Access Policy, s. f.), un repositorio digital gratuito en el que se archivaban artículos completos de acceso público, publicados dentro del ámbito biomédico y literatura de las ciencias biomédicas. Fue creado en 2014, dentro del Comité de Asignaciones de la Cámara de Representantes de Estados Unidos.

PubMed Central es una de las bases de datos más importantes dentro de los recursos desarrollados por el National Center for Biotechnology Information (Centro Nacional de Información Biotecnológica; en adelante, NCBI por sus siglas en inglés).

En realidad, es más que un repositorio de documentos, pues tiene un procedimiento para indexar y dar formato a los documentos con la finalidad de generar unos mejores metadatos, ontologías médicas e identificadores únicos que enriquecen los datos XML para cada uno de los artículos depositados. El contenido puede vincularse fácilmente con otros muchos contenidos.

PubMed Central es un archivo gratuito de artículos accesible a cualquiera, desde cualquier parte, a través de un simple buscador web. La totalidad del texto de los artículos de *PubMed Central* es de acceso libre para su lectura, aunque varían las condiciones de reutilización de la información.

Es obligatorio que todas las publicaciones estén disponibles en el repositorio en los doce meses siguientes a su publicación (Carroll, 2008). Por lo tanto, las publicaciones resultantes de las investigaciones tienen una licencia de OA que les es de aplicación y *PubMed Central* es el repositorio en el que los propios autores depositan sus publicaciones.

La política de acceso público del National Institute of Health en aplicación de la Ley de Presupuestos Estadounidense de 2008 (denominada Consolidated Appropriations Act), la División G del Título Segundo NIH, Sección 218 de la Ley Pública (PL por sus siglas en inglés) 110-161 establecía [traducción propia]:

El director del Instituto Nacional de Salud deberá exigir que todos los investigadores que reciban fondos del National Institute of Health presenten por sí mismos o hagan que se haga en su nombre una versión electrónica de su publicación final revisada por pares a la librería nacional de medicina *PubMed Central*, una vez que la publicación haya sido aceptada, para ponerla a disposición del público no más tarde de doce meses después de la fecha oficial de publicación. Sin embargo, el Instituto Nacional de Salud deberá implementar la política de acceso público de una forma consistente con el derecho de *copyright*. (Consolidated Appropriation Act, 2008, 345).

El Gobierno estadounidense financia la investigación con la expectativa de que las nuevas ideas y descubrimientos procedentes de la investigación impulsen la ciencia, estimulen la economía y mejoren las vidas y el bienestar de las personas.

Por ello, la Casa Blanca impulsa normas e iniciativas de agencias con la finalidad de que la investigación subvencionada con fondos públicos esté disponible y sea de acceso público. Promover el acceso a la investigación subvencionada con fondos públicos es prioritario en la medida en que dichas iniciativas mejoran el acceso general, proporcionan efectivas estrategias de archivo para estos recursos y aseguran el retorno de la inversión federal.

La necesidad de *open access* y de políticas de *public access* se ha puesto de manifiesto en repetidas ocasiones. La primera tentativa de *public access* fue la ley estadounidense de acceso público a la investigación federal (2012). El pilar de esta ley, como su propio nombre indica, es el acceso público a los resultados de las investigaciones realizadas con fondos federales. Este acceso público exige que se cumplan determinadas condiciones. Debía tratarse de proyectos realizados en el marco de once agencias y, además, su coste debería superar los cien millones de dólares.

Durante varias legislaturas esta ley no pudo aprobarse. Tenía grandes detractores, sobre todo los editores, contrarios a la libre publicación de los resultados. Después de presentarse en el Congreso en 2006, 2009 y 2012 y no aprobarse, los partidarios del *open access* decidieron proponer un texto nuevo: Fair Access to Science and Technology Research (FASTR) Act (2015), que podría traducirse como la ley de acceso equitativo a la ciencia y a la investigación tecnológica. Es una norma que persigue acelerar los descubrimientos científicos y espolear la innovación promoviendo el libre acceso a los artículos relacionados con la investigación científica financiada con fondos públicos, a través de Internet.

Aprobada en el Congreso estadounidense, esta ley fue publicada en 2013, y consolidaba la tendencia imparable a abrir cada vez más el acceso a los datos resultantes de la investigación tantas veces defendida desde la Casa Blanca.

En efecto, el Gobierno estadounidense subvenciona la investigación con la expectativa de que las nuevas ideas y descubrimientos de la investigación impulsen la ciencia, estimulen la economía y sirvan para mejorar la calidad de vida de los ciudadanos.

3.3.4.1. Las exigencias de la FASTR

El punto de arranque era la exigencia de conceder un acceso abierto a los artículos y datos revisados por otros expertos siempre que recabaran resultados de investigación financiada con fondos federales.

No obstante, esta exigencia, como decíamos, debe reunir varias condiciones previas. En primer lugar, solo afectaba a datos de once agencias concretas (Departamento de Agricultura, Departamento de Comercio, Departamento de Defensa, Departamento de Educación, Departamento de Energía, Departamento de Salud y Servicios Humanos, Departamento de Seguridad Interior, Departamento de Transporte, Agencia de Protección medioambiental, Administración Nacional del Espacio y Aeronáutica y la Fundación Nacional de Ciencias) siempre que la inversión realizada superase los cien millones de dólares.

La norma establecía unas directrices para las agencias. En efecto, al establecer la finalidad para la que se aprobaba, manifestó ser una ley cuya finalidad era que las agencias federales desarrollasen políticas de acceso abierto (Anon, 2015) en proyectos realizados por personal que trabajase para las agencias con fondos que dichas agencias les proporcionasen.

Entre las razones mencionadas como fundamento de la aprobación de esta norma, constaban las siguientes. En primer lugar, como ya se ha puesto de manifiesto, el hecho de que el Gobierno federal esté financiando proyectos con la finalidad de que

la ciencia avance y de mejorar la calidad de vida de los americanos y del mundo en general. En segundo lugar, también clave para este proceso, el impacto de Internet, que permite diseminar la información y ponerla a disposición de cualquiera fácilmente. En tercer lugar, el hecho de que Estados Unidos tenga interés en maximizar el impacto y la utilidad de la investigación que financia haciendo posible una amplia reutilización de los artículos que recojan los resultados de esta investigación financiada.

La norma no estableció las condiciones de la política de acceso público, pero dio un plazo de un año (que concluyó en febrero del 2014) para que cada una de las agencias desarrollase una política de acceso público. No obstante, sí que exigió que las normas de recogida y depósito de datos fueran lo más parecidas posibles. En primer lugar, era necesario solicitar tanto el depósito del artículo con los resultados como la revisión del artículo por otros expertos, y enviar a la agencia la última versión con todos los cambios cuando se cumplieran las siguientes condiciones: (i) que el editor consintiera poner la última versión y (ii) que se cumplieran los objetivos de la agencia federal en relación con la funcionalidad e interoperabilidad. Se garantiza facilitar el *acceso en línea gratuito* a los artículos revisados por otros expertos lo antes posible y a más tardar seis meses después de su publicación. Además, deberá darse acceso en condiciones que hagan posible la reutilización de la investigación incluyendo la aplicación de las últimas tecnologías existentes.

La norma exigió, asimismo, la elaboración de una bibliografía en línea de los artículos publicados, públicamente accesibles de conformidad con la política con un enlace que permitiese acceder al texto completo de los mismos. Debían cumplirse además tres condiciones: facilitar el acceso gratuito, garantizar la preservación de los datos y asegurar la interoperabilidad de los mismos (repositorios digitales estables).

3.3.4.2. Public access en la Administración Obama

En Estados Unidos, las distintas Administraciones Gobiernos, al llegar al poder, marcan las líneas que van a condicionar la política que desarrollarán en los años sucesivos.

En 2013, la Administración de Obama, recogiendo estas líneas maestras de *public access*, publicó en la página web de la Casa Blanca una carta vinculando el compromiso de esta Administración con el hecho de que los ciudadanos «se merecen un fácil acceso a los resultados de la investigación financiada con lo que han pagado con sus impuestos».

La Oficina Ejecutiva de la Presidencia publicó un memorándum de políticas el 22 de febrero de 2013 titulado *Increasing Access to the Results of Federally Funded Scientific*

Research en el que se fijaban los principios a los que quedaba sujeta esa Administración.

Cabe destacar en la introducción el compromiso de asegurar que los resultados directos de la investigación realizada con fondos federales se pusieran a disposición del público, de la industria y de la comunidad científica en la mayor medida posible y con el mínimo de restricciones posibles. Recogía, además, la vinculación con la American Competes Act de 2010, actualizada en 2015.

3.3.4.3. La Administración Trump

Parece que con la recién llegada Administración Trump va a haber cambios en relación con la gestión de los datos y la apertura de los mismos. Sin lugar a dudas, lo que trae esta Administración es incertidumbre en relación con muchos temas. No solo respecto a cómo lidiará esta Administración con los datos obtenidos de proyectos financiados con fondos públicos, sino también sobre si los resultados obtenidos pueden llegar a manipularse para justificar las políticas que se adopten (O'Brien, 2017).

De hecho, en febrero de 2017, la Administración Trump depuró la página web www.whitehouse.gov/briefing-room/disclosures, elaborada durante la Administración Obama. Es frecuente que el presidente entrante haga modificaciones, pero lo relevante en este caso es que desaparecieron muchos datos y conjuntos de datos (O'Brien, 2017). Muchos de los puestos creados en la Administración Obama para llevar a cabo toda esta gestión han desaparecido o están vacantes hoy con la Administración Trump.

La otra gran pregunta aún sin respuesta es la cantidad de subvenciones que se concederán. En efecto, tan grave es limitar el acceso a los datos como disminuir las subvenciones (Wojick, 2016); de hecho, una de las grandes incógnitas es la reducción de la distribución de subvenciones. Todavía es pronto para hablar de las líneas en las que trabajará la Administración Trump, aunque lo que sí queda claro es que ha marcado una ruptura con la Administración Obama.

En julio de este año tuvo lugar una mesa redonda sobre *open data* con la Casa Blanca (Howard, 2017) en la cual la Administración Trump puso de manifiesto su voluntad de apoyar la continuación de la divulgación pública de datos públicos en línea. De hecho, la Office of Management and Budget (OMB, Oficina de Gestión y Presupuestos) defiende estar doblando la inversión en ello mediante la modernización de la tecnología.

La OMB puso su atención en el *open data* y el hecho de que la Administración tiene localizadas muchas oportunidades para usar *open data* con la finalidad de mejorar el

Gobierno. La Casa Blanca está centrada en una serie de cuestiones en las que los datos abiertos son más relevantes como se muestra en la siguiente figura.

DATA.GOV Open Data Fuels Economic Growth

U.S. Federal Open Data is a strategic national resource. The value of Open Data to the United States has been estimated at hundreds of billions of dollars. American businesses use it to improve their operations, target their marketing, and develop new products and services.

Have a case study example? Share your story at data.gov/impact.

	Finance <ul style="list-style-type: none"> • Evaluate personal loan applications. • Assess small businesses seeking financing. 	<ul style="list-style-type: none"> • Inform investment decisions in real estate, currencies, commodities, and more. • Develop research on national and global economic and financial trends.
	Healthcare <ul style="list-style-type: none"> • Tailor treatments to individual needs with Precision Medicine. • Develop information products for patients, doctors, and hospitals. 	<ul style="list-style-type: none"> • Develop and improve pharmaceutical drugs and medical devices. • Predict, monitor, and respond to disease outbreaks.
	Consumer & Retail <ul style="list-style-type: none"> • Improve inventory management based on weather and other factors. • Better manage their supply chains. 	<ul style="list-style-type: none"> • Create targeted marketing strategies with consumer segmentation models. • Plan their investments in market expansion, diversification, and acquisition.
	Transportation & Trade <ul style="list-style-type: none"> • Monitor vehicle safety and improve safety features. • Calculate risks to inform pricing of insurance products. 	<ul style="list-style-type: none"> • Optimize routes, supply chains and logistics across all modes of transportation. • Support R&D and product innovation, for instance to develop autonomous vehicles.
	Geospatial <ul style="list-style-type: none"> • Build mobile applications, GPS navigation tools, and mapping platforms. 	<ul style="list-style-type: none"> • Provide geospatial analytics and visualizations for understanding data. • Find trends in population growth, traffic patterns, and public health.

Data.gov is the home of the U.S. Government's open data where you will find data, tools, and resources to develop web and mobile applications, design data visualizations, conduct research, fuel private-sector business models, and more.

ROUNDTABLE ON OPEN DATA FOR ECONOMIC GROWTH

Figura 38. Notes from a roundtable on open data at the White House. Fuente: sunlightfoundation.com/2017/07/26/white-house-open-data-roundtable.

Parece que, poco a poco, se va aclarando el posicionamiento de la nueva Administración en lo que respecta a los datos abiertos. Habrá que ver el impacto que tiene en la ciencia.

Adicionalmente, durante el mandato Trump desaparece la neutralidad en la red. Es el principio por el cual los proveedores de servicios de Internet y los Gobiernos que la regulan deben tratar todo tráfico de datos que transita por la red de igual forma indiscriminada, sin cobrar a los usuarios una tarifa dependiendo del contenido, página web, plataforma o aplicación a la que accedan. Tampoco según el tipo de equipamiento, dispositivo o método de comunicación que utilizan para el acceso. En definitiva, la neutralidad de la red evita que el proveedor contratado cargue tarifas adicionales por visitar alguna web a su cuenta por el servicio prestado. Si este principio es vulnerado, se perderá no solo la libertad de ver lo que cada uno quiera en la red, sino que, además, cobrarán más por acceder a ello.

En el fondo se trata de un problema de precios: consiste en determinar si los operadores de telecomunicaciones y los proveedores de acceso a Internet deben cobrar o no a los consumidores solo una vez por el acceso a Internet, sin favorecer a un proveedor de contenidos sobre otro, ni imponiendo a los proveedores y plataformas de contenidos precios adicionales por enviar información a través de sus redes a los internautas finales⁴⁰.

Este tema es tan reciente que no es posible evaluar todavía el impacto que tendrá en la red a corto, medio y largo plazo. Por el momento, el debate que se ha producido es más un debate político que real. Si bien las grandes tecnológicas como Google, Facebook o Amazon temen que pueda utilizarse para frenar la innovación que están llevando, aún no se han producido grandes cambios.

3.3.4.4. ¿Una excepción al *public access*?: la University and Small Business Patent Procedures Act más conocida con el nombre de ley Bayh-Dole

Durante los años setenta surgió un malestar tecnológico cuyo origen estaba en el hecho de que la investigación sufragada con fondos públicos fuera propiedad de la agencia que hubiera realizado la inversión.

La inversión del Gobierno resultaba ineficiente para licenciar las patentes registradas a explotadores privados. Como consecuencia de ello, los hallazgos resultantes de investigación financiada con fondos federales rara vez llegan a convertirse en productos con uso comercial (O'Brien, 2013).

⁴⁰ Recuperado de: https://es.wikipedia.org/wiki/Neutralidad_de_red. (Consultado por última vez el 30 de marzo de 2018).

Antes de 1981, ni los científicos ni las universidades tenían la propiedad de ninguna invención creada con fondos estatales. La investigación financiada con fondos públicos pertenecía al público. En Estados Unidos, históricamente, los monopolios están denostados y dado que la patente es un monopolio es, por lo tanto, un mal necesario (Valoir, 2000). Tanto el presidente Kennedy durante su mandato y a principios de los años setenta como el presidente Nixon en el memorando en el que se establecieron los principios aplicables a las patentes (White House, 1971) trataron de corregir esto.

Estas preocupaciones constituyeron los antecedentes de la aprobación de la reforma operada por la ley de Procedimientos de Patentes para Universidades y Pequeños Negocios, más conocida con el nombre de ley Bayh-Dole, que fue aprobada en 1980. El objetivo principal de esta modificación era desarrollar la innovación estadounidense y, al mismo tiempo, promover la participación de las pequeñas empresas, así como lograr que las invenciones estuvieran disponibles para el público en general. Al mismo tiempo, se aseguraban de que el Gobierno obtuviera suficientes derechos en las invenciones financiadas con fondos federales para poder satisfacer sus propias necesidades y las del público en general (Valoir, 2000).

Por lo tanto, la reforma del Congreso tuvo una triple finalidad (O'Brien, 2013):

1. Asegurar la efectiva transferencia y el desarrollo comercial de las patentes que tienen su origen en la investigación financiada por el Estado.
2. Revigorizar la industria estadounidense proporcionándole una fresca inyección de nuevas ideas.
3. Asegurarse de que lo que se descubra en el marco de la investigación financiada por Estados Unidos lo desarrollan empresas estadounidenses y no competidores extranjeros.

De alguna forma, esta ley pretendía permitir a los pequeños negocios y a las entidades académicas que recibiesen fondos federales para investigación, que pudieran ser los titulares de las patentes resultantes y que se pudieran transmitir derechos exclusivos sobre estas patentes a empresas privadas que pudieran explotarlas.

Si bien la principal finalidad de esta norma era promover la inversión privada y la innovación, el legislador tenía también en mente proteger al público en general del posible perjuicio que pudiese resultar de que una empresa no hiciera uso o hiciera un uso no razonable de una invención financiada con fondos públicos. En efecto, con esta finalidad la ley recogió dos reservas significativas.

En primer lugar, el Gobierno federal retuvo una licencia no exclusiva, no transferible, gratuita para usar la patente para y en nombre de Estados Unidos. En

segundo lugar, la agencia federal financiadora, siempre que se reuniesen ciertas condiciones, podía ejercitar sus derechos de *march-in*⁴¹ y obligar al titular de la patente a conceder una licencia a un solicitante responsable. Desde entonces, la agencia federal tiene la posibilidad de ejercitar ese derecho, directamente o a instancia de parte.

Los derechos llamados *march-in* tienen una larga historia. Su nacimiento se remonta a 1947 (Raubitschek y Latker, 2005) en las recomendaciones que el abogado general hacía en su informe al presidente Roosevelt y que también quedaban recogidos en los memoranda de los presidentes Kennedy y Nixon.

Aunque existentes antes de la ley Bayh-Dole, no se han ejercitado nunca si bien según Donald R. Runner, vicepresidente de la American Patent Law Association, en ningún caso concreto debió ejercitarse. Una de las consideraciones principales de la ley Bayh-Dole fue la de encontrar el equilibrio entre el beneficio público y el beneficio privado (Valoir, 2000).

La dificultad procedimental que tenía el ejercicio de los derechos de *march-in* hizo que una parte de la doctrina los considerase, en última instancia, ineficientes para situaciones de emergencia. Además, el ejercicio de derechos de *march-in* exigía que la producción de los productos a los que se aplicase la «expropiación» que permiten los derechos de *march-in* productos fuese sustancialmente realizada en Estados Unidos. A pesar de las previsiones de la ley Bayh-Dole, la realidad es que en Estados Unidos, nunca se han ejercitado los derechos de *march-in* a pesar de haberse instado, una vez aprobada la ley Bayh-Dole, varias veces ante el Congreso.

Esta falta de ejercicio de los derechos de *march-in* ha sido objeto de muchas discusiones en la doctrina. Se ha solicitado su ejecución, al menos, en tres supuestos: uno en 1997, en el caso Cell Pro (O'Brien, 2013; Eberle, 1999); otro, en 2004, respecto a Norvir y Xalatan (O'Brien, 2013) y el más reciente en 2009, con el caso Fabrazyme (O'Brien, 2013), que servirá de ejemplo para ilustrar y comprender de qué manera funcionan los derechos de *march-in*.

Fabry era una enfermedad muy rara y dolorosa que, en última instancia, podía ser letal. Con financiación del National Institute of Health, se descubrió una encima para tratar esta enfermedad, cuya patente se cedió a Genzyme para que hiciera todo

⁴¹ No se ha encontrado una buena traducción para el término, por lo que se procede a explicarlo. La posibilidad de *march-in* supone una reserva de la agencia para «paralizar» de alguna forma la licencia concedida y conceder una licencia a otro solicitante para proteger a los estadounidenses. Tiene su justificación en que las patentes tienen su origen en financiación pública y, por lo tanto, por razones de salud o de protección de los ciudadanos estadounidenses, las agencias financiadoras deberán actuar y atribuir la licencia de la patente que no esté siendo explotada o que no le esté siendo de forma razonable.

el desarrollo de la molécula. Como consecuencia de un fallo en la producción se produjo una falta de producto en el mercado estadounidense.

A resultas de este problema de producción, Genzyme, además de pagar 175 millones de dólares, se vio obligada a restringir la producción, por lo que se pasó de suministrar a los pacientes estadounidenses una dosis al mes de las dos que estaban recibiendo antes de las restricciones en la producción. Adicionalmente, el medicamento no podía prescribirse a los pacientes a los que se diagnosticara la enfermedad. Aunque un grupo de pacientes requirieron en 2010 al National Institute of Health para que ejercitara los derechos de *march-in*, este se negó sobre la base de las declaraciones de Genzyme: manifestó que esperaba alcanzar el mismo nivel de producción durante la primera mitad de 2011. Las complicaciones en la producción continuaron y Genzyme tuvo que volver a retrasar la fecha.

Como consecuencia de las reducciones de dosis, la Agencia Europea del Medicamento (EMA, por las siglas en inglés de European Medicines Agency) detectó una multiplicación de los efectos adversos en los pacientes y, por ende, exigió a Genzyme que mandara cantidad suficiente para que todos los pacientes tuvieran acceso a la dosis total. Nuevamente, los pacientes recurrieron al National Institute of Health para que ejercitara sus derechos de *march-in* y para que prohibiera la exportación del producto a Europa con la finalidad de que los pacientes estadounidenses tuvieran acceso a la medicación. La recuperación de los niveles de producción normales se demoró más de un año.

Sin embargo, el National Institute of Health se justificó por no haber ejercitado los derechos de *march-in* basándose en las siguientes consideraciones:

1. Puso de manifiesto que el hecho de dar una licencia para producir Fabrazyme a un tercero no iba a suponer ninguna ventaja a corto plazo, dado que la obtención de la aprobación de la FDA para comercializar el medicamento tardaba mucho tiempo.
2. Además, manifestó que ningún tercero había solicitado a la FDA la aprobación de ningún producto como alternativa al Fabrazyme. Este argumento no parecía muy consistente, puesto que cualquier intento de hacerlo hubiera debido enfrentarse a una posible reclamación por infracción de patente.
3. Por último, declaró que Genzyme se había comprometido a resolver el problema en seis meses y que había estado trabajando para conseguirlo de forma diligente y de buena fe.

Esta actuación del National Institute of Health fue muy cuestionada, pues parecía contraria a una de las primeras finalidades de esta ley, es decir, proteger al público de

la falta de uso o del uso irracional de las invenciones financiadas con fondos públicos. Una gran parte de la doctrina considera que el hecho de que ninguna agencia haya ejercitado los derechos de *march-in* reduce las posibilidades de que alguna empresa trate de presentarse como alternativa en caso de riesgo para la población estadounidense, puesto que sería fácil que incurrieran en infracción de una patente concedida de forma válida. Se preguntaron, incluso, si el hecho de ejercitar los derechos de *march-in* no incrementaría las posibilidades de que pudieran aparecer alternativas en caso de necesidad.

Por lo tanto, si bien el Congreso pretende conseguir un equilibrio entre los intereses privados y los públicos reteniendo los derechos de *march-in* sobre la investigación financiada con fondos federales, los casos que se han ido planteando conducen a pensar que la ley Bayh-Dole ha logrado el primero de los objetivos (asegurar la efectiva transferencia y el desarrollo comercial de las patentes que tienen su origen en la investigación financiada por el Estado permitiendo que los pequeños negocios y las entidades académicas que pudieran transmitir derechos exclusivos sobre las patentes a empresas privadas para que estas las explotaran), pero no ha ocurrido lo mismo con el segundo (proteger al público en general del posible perjuicio que pudiera resultar del hecho que una empresa no hiciera uso o hiciera un uso no razonable de una invención financiada con fondos públicos). En efecto, el caso *Fabrazyme*, demuestra, dicen, que no ha conseguido proteger la salud pública y la seguridad de la falta de uso o uso no razonable de invenciones financiadas con los impuestos de los contribuyentes estadounidenses (O'Brien, 2013).

A pesar de todas las discusiones aparecidas en Estados Unidos en relación con la ley Bayh-Dole, lo cierto es que este modelo sentó un precedente y ha tratado de emularse en otros mercados (Eisenberg, 1996). La ley Bayh-Dole supuso una proliferación del aprovechamiento real y comercial de la investigación financiada con fondos públicos para los estadounidenses, aunque ello no ha evitado que hayan ido surgiendo voces discrepantes. En efecto, de alguna forma, la posibilidad de que la empresa privada patente descubrimientos realizados con fondos públicos podría llegar a limitar los avances de la ciencia, por lo que es necesario encontrar un equilibrio.

Después de más de veinticinco años de vigencia, algunos autores se han planteado si esta ley —y la flexibilidad que propuso— es la adecuada o es necesario realizar una modificación de la regulación y aprobar una política pública uniforme. Un punto de flexibilidad de la regulación ha permitido dar respuesta a una multiplicidad de problemáticas y, al mismo tiempo, ha generado algunos desalineamientos de las respuestas (Boettiger y Bennet, 2006).

Todo este análisis lleva a analizar, ahora, cómo todo esto ha impactado en el intercambio de datos. Como puede verse, tanto en Estados Unidos como en

Europa existe un creciente enfoque en lograr mejorar el intercambio de datos, motor de la revolución digital actual. Sin embargo, procede analizar antes toda la problemática previa que ha aparecido.

4. EL *DATA SHARING* O INTERCAMBIO DE DATOS

Estos dos términos se usarán indistintamente a partir de ahora. Después de un análisis en abstracto de ambos modelos, procede analizar todas las cuestiones de intercambio de datos en dos instituciones globales, enfocadas ambas, principalmente, a compartir datos de investigación o científicos.

Los datos se han convertido en la infraestructura de la ciencia. Ya en 2010, se puso de manifiesto que «Science is becoming data-intensive and collaborative» (Seidel, 2010),⁴² razón por la cual compartir datos se hace cada vez más clave. El intercambio de datos consiste en poner los datos a disposición de los demás para que puedan usarlos (Michener, 2015). Los ecologistas cada vez generan más datos, que es fundamental poner en común para el desarrollo del conocimiento. Esto es clave, hoy, a la hora de gestionar temas de biodiversidad, en la medida en que permite abordar cuestiones de alto nivel (Enke *et al.*, 2012; Gaikwad y Chavan, 2006).

Compartir información es esencial para los científicos (Clearwater, 2010). Un estudio de la Research Information Network y de la British Library señala que «la mayor parte de los científicos pasa mucho de su tiempo buscando y ordenando información» (Research Information Network y British Library, 2009).

Por lo tanto, el intercambio de datos es la práctica de poner los datos a disposición de otros o de reutilizarlos para subsiguientes análisis, e incluye también cuestiones de almacenamiento de datos. Las bases de datos proporcionan la implementación técnica de los almacenamientos de la información, mientras que las organizaciones o instituciones son las que proporcionan los repositorios (Enke *et al.*, 2012). Existe, por consiguiente, una correlación entre el intercambio de datos y su reutilización. Casi puede decirse que no se entiende uno sin el otro. A fin de cuentas, la finalidad de poder intercambiar datos es la reutilización de los mismos.

Idealmente, cada uno de los descubrimientos científicos debería ponerse a disposición de todos (Merton, 1973). Los beneficios de este intercambio son incuestionables (Fienberg *et al.*, 1985; Vickers, 2006). Uno de los beneficios fundamentales es el mayor retorno de la inversión obtenida (Torres-Salinas *et al.*, 2012) y también es clave para evaluar la calidad de los datos (Costello *et al.*, 2013). Requiere ponerlos a disposición de los demás y no solo publicarlos.

⁴² Traducción propia: «La ciencia se está haciendo intensiva en datos y colaborativa».

El hecho de que los datos estén en Internet no es suficiente para que puedan usarse. En la actualidad existe toda una corriente hacia la que van a necesitar ir convergiendo los dos tipos de regulaciones, la existente en Europa y la de Estados Unidos, aunque partan de puntos diferentes: compartir datos, reducir al máximo las restricciones en el uso de los datos y permitir su reutilización. Cada vez más se tiende hacia abrir el conocimiento obtenido con fondos públicos (Comisión Europea, National Science Foundation). Es más una tendencia de políticas a alto nivel, de principios, a la que tendrán que ir acercándose las normas de los distintos países.

A pesar de la multiplicidad de intentos para que los científicos pongan los datos a disposición de todos, esto no ocurre en la mayor parte de los casos de biodiversidad y medioambiente (Costello, 2009). No se ha dado respuesta a las preocupaciones de los científicos y las iniciativas para motivarlos han resultado inadecuadas.

Los científicos esgrimen toda una serie de razones para no poner los datos a disposición de todos (Costello, 2009):

1. *Copiarán mi trabajo de la web y harán plagio.*

La facilidad de copiar los datos en línea hace que muchos científicos estén incómodos con las publicaciones en línea. Sin embargo, publicando en línea la fuente será mucho más visible y el plagio más fácil de detectar.

2. *¿Dónde puedo publicar mis datos? Las revistas no publican datos primarios.*

Cada vez un mayor número de revistas en línea permiten publicar apéndices con los artículos donde pueden incluirse los datos primarios. Además, hay centros de datos que archivan y ponen a disposición de los demás cada vez un rango más elevado de datos de biodiversidad y medioambientales.

3. *Son mis datos,⁴³ ¿por qué tengo que ponerlos a disposición de nadie?*

En la mayor parte de los casos, los científicos son los que custodian los datos, aunque la propiedad de los mismos corresponde al empleador o a la organización; en el supuesto de que la recopilación de datos estuviera subvencionada por fondos públicos, los datos serán propiedad de la sociedad. Desgraciadamente, algunos científicos piensan que son los propietarios de los datos a pesar de que trabajan financiados con fondos públicos. En realidad, son administradores de datos públicos y, así, son los responsables de difundirlos.

⁴³ *It's my data*. Este es el conocido *it's my data syndrome* al que se hará referencia más adelante. Básicamente, los científicos históricamente tenían sensación de propiedad de todas las investigaciones que realizaban. Esto es altamente discutido hoy.

4. *Los datos que usé no son míos y no tengo permiso para publicarlos.*

Si el propietario de los datos permite a un científico usarlos, es muy posible que le autorice a publicarlos. Por lo tanto, la cuestión que se plantea es cuándo se solicita o no el permiso de publicar. Tal vez dado que no ven un beneficio personal en hacerlo, se tardará algo más de tiempo en actividades más beneficiosas o consideren que no tienen ninguna obligación ética de facilitar el acceso a los datos a otros.

5. *Si divulgo datos puede ser que alguien me gane y termine produciendo información para otro. Todavía no he terminado de analizar los datos y puede que necesite analizarlos con más profundidad.*

En muchos casos, es aceptable retrasar la publicación de datos hasta que se hayan analizado y se haya publicado la síntesis. Es importante que los datos se publiquen antes o justo después de hacerse la primera publicación para que puedan citarse. Si después de doce meses no se ha hecho ninguna publicación, es recomendable publicar los datos, para obtener al menos cierto reconocimiento.

6. *Alguien puede usar mis datos y beneficiarse de dicho uso o, peor aún, puede hacerlo una consultoría o una organización comercial.*

Los científicos quieren que el mundo saque provecho de los resultados de su investigación y, por eso, los publican en medios escritos; lo mismo es aplicable a los datos. Después de publicados, los autores dejan de tener el control sobre quien use sus descubrimientos y con qué fines, así que intentarlo es un gasto energético que no tiene sentido. Con carácter general, los científicos agradecen cualquier uso y cita de su investigación en medios escritos, incluso aunque lo hagan competidores. Entonces, ¿por qué no hacerlo también con los datos?

7. *El editor puede beneficiarse.*

Con carácter general, los científicos renuncian al *copyright* en beneficio de los editores, aunque la mayor parte de la investigación está financiada con fondos públicos y el poder beneficiarse de los resultados generaría más financiación.

8. *Me da miedo que los datos puedan usarse con una finalidad inadecuada.*

Los conjuntos de datos solo deben publicarse con los suficientes metadatos o con la documentación que describa cómo se han recabado y las limitaciones que tienen. Por lo tanto, será responsabilidad de los usuarios leerlos e interpretarlos bien.

9. *No tengo los conocimientos necesarios para publicar los datos en Internet.*

Cualquier persona que tenga los conocimientos necesarios para gestionar datos en tablas y hojas de Excel tiene las habilidades necesarias para entregar los datos en un formato estándar que las organizaciones puedan publicar en Internet.

10. *Los derechos de propiedad intelectual relacionados con los datos y las bases de datos varían según los países.*

Otra de las preocupaciones es la pérdida de los derechos de propiedad intelectual, de la propiedad, la autoría o del control de los datos, al ponerlos a disposición de todos. Este promueve las comparaciones entre las distintas normativas de derechos de propiedad intelectual en lo que respecta a los datos, las bases de datos y las publicaciones. La complejidad resultante desincentiva el intercambio de datos. Sin embargo, parece que esto no es aplicable si se considera la publicación de datos como la publicación por escrito. Por lo tanto, estos condicionantes no limitan la publicación en medios escritos. De hecho, los modelos de publicación de datos en línea de algunas organizaciones permiten atribuir a los custodios de los datos mayor control sobre la publicación de los mismos que los medios convencionales.

11. *No voy a obtener el reconocimiento merecido como consecuencia de la creación de los datos.*

La publicación es la forma más clara para que los científicos puedan obtener crédito público de sus ideas y de su trabajo. Asimismo, la mejor manera de obtener reconocimiento por haber recopilado datos útiles es publicándolos.

La ciencia obtendrá mayores beneficios en el supuesto de que los datos se publiquen en abierto, como se hace por ejemplo con las licencias de Creative Commons. Un requerimiento común para los usuarios es la atribución de la fuente, incluyendo sus autores y editores. Si los usuarios de los datos citan claramente sus fuentes, igual que lo harían en el caso de publicaciones en medios escritos, añaden credibilidad a los datos utilizados de esta manera. Si no lo hacen, pueden incurrir en plagio.

12. *Otras razones.*

Al igual que ocurre en las publicaciones en medios escritos, la publicación de datos puede poner al descubierto problemas con los datos y, por ende, contribuir a la mejora de su calidad. Algunas veces los datos se organizan con una determinada idiosincrasia que dificulta que otras personas puedan analizarlos. Las filas y columnas pueden no estar adecuadamente categorizadas, consistentemente formateadas ni contar con la suficiente descripción para que otros lo entiendan. Los autores pueden temer que se haga un uso selectivo de los datos o que existan errores de análisis. No obstante, exponer los datos al análisis independiente de terceros es un beneficio para la ciencia, puesto que

puede suponer un afinamiento o llegarse a unas conclusiones distintas de aquellas a las que llegó el autor.

Todas estas cuestiones siguen, de alguna manera y a distintos niveles, presentes. Las políticas expresas de las revistas que requieren intercambio de datos no llevan a los autores a poner sus datos a disposición de investigadores independientes (Savage y Vickers, 2009).

Todavía queda un largo camino por recorrer; sin embargo, pueden verse algunos avances. Un estudio realizado en 2011 en relación con la percepción que los científicos tienen del hecho de compartir datos lo pusieron de manifiesto Tenopir *et al.* (2011). En 2014, se realiza un nuevo estudio realizado como seguimiento al de 2011 (Tenopir, 2015); las conclusiones extraídas del mismo merecen una reflexión.

Parecía existir una mayor conciencia de la necesidad de compartir los datos y, curiosamente, se observó que el nivel de intercambio de datos era mucho más frecuente en científicos de más edad, ya consolidados. Los jóvenes, aunque conscientes de la necesidad de compartir, se mostraron menos dispuestos a hacerlo, tal vez más preocupados por las publicaciones. Todo esto quiere decir que, si bien una multiplicidad de organizaciones están trabajando a alto nivel en principios y conceptos de apertura en la ciencia, todavía existe un gran desajuste entre esas políticas de alto nivel impulsadas por la Comisión Europea, la National Science Foundation de Estados Unidos, la National Science Foundation de Australia, etc., y lo que realmente está ocurriendo con los científicos que están trabajando en los proyectos.

Los científicos que investigan la biodiversidad están dispuestos a compartir e intercambiar sus datos, pero no acaban de hacerlo por razones de distinta índole: tecnológicas (falta de las bases de datos adecuadas y de los mecanismos de cita) y sociológicas (tiempo, financiación, sostenibilidad) (Enke *et al.*, 2012). La mejora de los intercambios de datos requiere una mejora de las infraestructuras técnicas.

Identificada la necesidad del intercambio de datos, queda por abordar cómo contribuir, desde el punto de vista de eliminación de barreras jurídicas, a este intercambio de datos. En efecto, si el intercambio de datos es el fin, la eliminación de las barreras es el medio. Tal como puede verse en el análisis analizado antes (Costello, 2009), la percepción de los científicos de las dificultades jurídicas y la variedad de regulaciones está sirviendo de barrera o, cuando menos, de excusa para no alcanzar el nivel de intercambio de datos deseable.

Una serie de organizaciones están adoptando políticas de intercambio de datos para facilitar una estructura donde puedan hacerse y una serie de principios que lo hagan posible.

4.1. El Sistema de Sistemas Global de Observación de la Tierra

4.1.1. Los antecedentes del GEOSS: GEO

En la Primera Cumbre celebrada en 2003 en Washington DC, representantes de treinta y tres países, la Comisión Europea y más de veinte organizaciones internacionales pusieron de manifiesto la necesidad de que existiese un sistema de observación de la Tierra, integral, coordinado y sostenible.

En febrero de 2005, en la Tercera Cumbre sobre la Observación de la Tierra (Earthzine, s. f.) celebrada en Bruselas y al final de un proceso que comenzó en la Primera Cumbre de la Observación de la Tierra en Washington, se creó el Grupo de Observación de la Tierra (GEO, por las siglas en inglés de Global Earth Observation) y, dentro de ese marco, se empezó a fraguar el Sistema Mundial de Sistemas de Observación de la Tierra GEOSS (Group on Earth Observations, s. f.) (por las siglas en inglés de Global Earth Observation System of Systems), como respuesta a las llamadas a la acción de la Cumbre Mundial de Desarrollo Sostenible de 2002 celebrada en Johannesburgo (Uhlir *et al.*, 2009) y del Grupo de los Ocho (G8), formado por los líderes de los países industrializados.

En estas reuniones de alto nivel, se reconoció que la colaboración internacional era fundamental para explotar el potencial de crecimiento de las observaciones de la Tierra y para servir de soporte a la adopción de decisiones en un mundo cada vez con mayor complejidad y estrés medioambiental. En esta cumbre, se adoptó el compromiso político de ir desarrollando GEOSS (GEO Secretariat, 2005).

El GEO nació como un grupo *ad hoc*, intergubernamental y copresidido por la Comisión Europea, Japón, Sudáfrica y Estados Unidos, que se encargó de la tarea de desarrollar un plan a diez años, que debía ser presentado en febrero de 2005.

Hoy GEO (2008) es una organización integrada por más de ciento dos países miembros y más de ciento tres organizaciones. Trabaja para mejorar la coordinación de los datos existentes en materia de observación de la Tierra, desarrollar nuevas observaciones y promover la generación de productos de observación de la Tierra. estaba enfocada a dar acceso a los datos de observación de la Tierra en nueva áreas prioritarias: desastres naturales y provocados por el hombre, fuentes medioambientales de riesgos para la salud, gestión de la energía, cambio climático y el impacto que tiene, recursos de agua potable, previsiones meteorológicas, gestión de ecosistemas, agricultura sostenible y conservación de la biodiversidad.

Para el GEO fue fundamental la construcción del GEOSS sobre la base de un plan a diez años para implementarlo; por eso debía supervisar la articulación de los mecanismos para lograr estos objetivos. Se crearon cinco subgrupos técnicos y una pequeña Secretaría que elaboró un documento marco adoptado para su ejecución en

una cumbre celebrada en Tokio en 2004. En este documento marco quedaron definidos el alcance y el objetivo del GEOSS (GEO Secretariat, 2005).

En los últimos treinta años, la disponibilidad de datos geoespaciales ha tenido un crecimiento sin precedente con la evolución de la tecnología (Philips *et al.*, 2012). Tal como ya se habló en el punto relacionado con la revolución digital, uno de los retos que hay que abordar hoy es darle sentido a la ingente cantidad de para que se convierta en conocimiento comprensible (Gore, 1998). Muchos usuarios consideran difícil y costoso acceder a los datos necesarios. También puede ocurrir que los datos estén recogidos de una forma difícil de interpretar o carezcan de la calidad necesaria.

A todo esto hay que añadir que los intercambios de datos entre agencias y países son insuficientes, en parte, por la incompatibilidad de las políticas de datos. Existe, también, una involucración inadecuada de los usuarios de datos en la especificación de los requerimientos de información. En ocasiones los retrasos a la hora de acceder a los datos impiden usar la información en el momento oportuno en el que podrían salvar vidas o minimizar pérdidas en propiedades. La generación y disseminación de productos que usan grandes volúmenes de datos muchas veces queda muy por detrás de las observaciones recogidas.

La cobertura espacial y temporal no está optimizada, lo que deja muchas partes del globo sin la suficiente información y disminuye de esta manera la eficacia de los sistemas de muestreo de las regiones en las que existen observaciones adecuadas. Puede ocurrir que no puedan combinarse observaciones de la misma variable en distintos lugares por parte de distintas agencias, porque los métodos usados para medir sean diferentes, no sigan estándares acordados o no estén correctamente intercalados, o porque la resolución en el tiempo y en el espacio o en las estructuras de datos en los que se almacenan las observaciones sean claramente incompatibles.

Los esfuerzos de observación se van repitiendo, innecesariamente, como consecuencia tanto de la falta de coordinación como de la imposibilidad de que un cierto número de usuarios diferente puedan hacer uso de una observación. Los sistemas de observación y redes están diseñados de forma separada en distintos ámbitos. Además, no se están realizando observaciones sostenibles sistemáticas y operacionales de muchos tópicos de vital importancia para la sociedad.

Algunos sistemas existentes no están operando de acuerdo con su capacidad (GEO Secretariat, 2005). Ejemplo de ello es el tsunami de 26 de diciembre de 2004, donde se puso de manifiesto la necesidad de que exista un sistema internacional de observación y predicción coordinado con servicios nacionales de emergencia para que el público reciba aviso a tiempo. La inexistencia de estos sistemas hizo que murieran cerca de doscientas sesenta mil personas (GEO Secretariat, 2005).

4.1.2. El GEOSS

La lógica que subyacía en el GEOSS era diferente: se requería poder utilizar las imágenes obtenidas desde los satélites para completar, complementar otras observaciones entre las que se incluían las de biodiversidad. En efecto, todas las observaciones estaban ligadas.

Nació como un sistema de sistemas (Christian, 2005), es decir, tenía vocación de integrar todos los observatorios, herramientas, sistemas, etc., existentes en la Tierra de manera que las decisiones se tomaran con el máximo de información posible. Perseguía lograr que se compartiesen las observaciones y productos con todo el sistema y que, de esta manera, se adoptasen medidas comparables y comprensibles apoyadas en estándares comunes.

La conceptualización de un sistema de sistemas fue fundamental para abordar todas las cuestiones derivadas de la interacción de sistemas operativos heterogéneos e independientes para la consecución de un único propósito (De Laurentis *et al.*, 2007). Ya en 2006 Latenbacher manifestó que este sistema representaría un cambio revolucionario para la Tierra y para la forma de comprenderla en los próximos años.

Por consiguiente, su objetivo era hacer posible la interconexión de la información de sistemas que entonces, y aún hoy, no estaban conectados para obtener una visión que recogiera las necesidades de los usuarios. Sin lugar a dudas era una herramienta digital con vocación de optimizar el conocimiento, de acercarlo, conservarlo, facilitar el acceso, etc., en la era del conocimiento actual.

Este sistema debía tratar de identificar gaps y duplicidades innecesarias con la finalidad de suprimirlas. Proporcionaba, por lo tanto, un marco conceptual general que permita construir un sistema integrado de observación de la Tierra

Se creó así una estructura *global* en dos sentidos. Global con vocación de convertirse en mundial y, por ende, abierta a todos los países y organizaciones que quisieran formar parte de la misma y global también en el sentido de la escala a la que debía hacerse, pues no afectaba a países individuales, sino al mundo en su totalidad. Para gestionar problemas interplanetarios, era necesario tener información interplanetaria (Stone, 2010).

Este sistema tiene como finalidad *la observación de la Tierra* y, por tanto, facilitar el acceso a las observaciones directas y a los productos. De igual manera, debe realizar procesos para que el sistema funcione como pueden ser el mantenimiento de los datos, el intercambio de estándares, etc. Las observaciones del GEOSS incluirán observaciones de todos los lugares de la Tierra, incluidos los que no pertenecen a ningún país como la Antártida o los océanos, así como del espacio.

El GEOSS, con esfuerzos de cooperación y partiendo de las estructuras ya existentes, pretendía crear una red flexible de proveedores de contenidos que permitieran acceder a un extraordinario rango de datos y de información desde su oficina a aquellos que tenían que tomar decisiones (Nativi *et al.*, 2015).

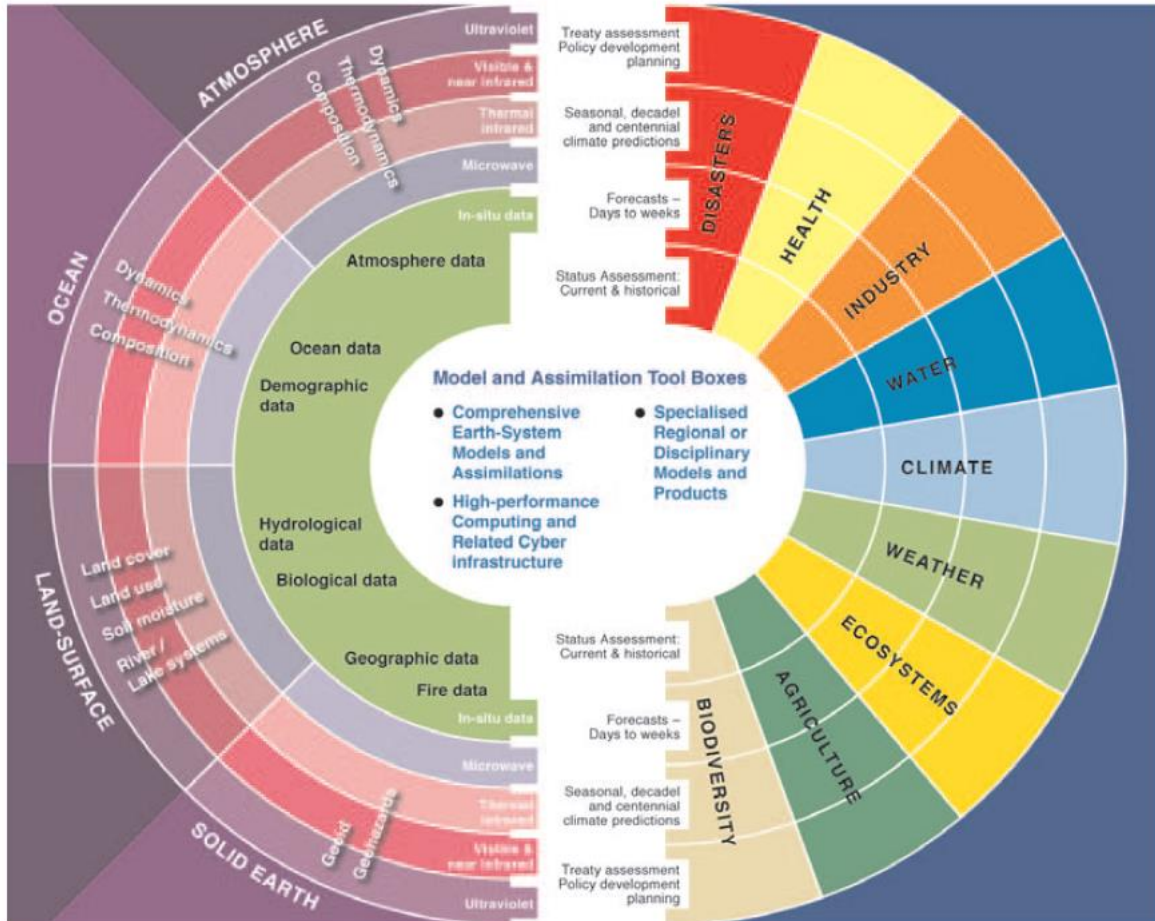


Figura 39. Basado en el Plan de Implementación de GEOSS, esquema de la transformación de las observaciones en información predictiva y de situación actual. En la parte derecha: productos de un sistema de pronóstico de sistemas terrestres y modelos especializados asociados organizados en las categorías de beneficios socioeconómicos de GEOSS, estratificados por el tiempo de entrega en el que requieren los entregables (evaluaciones de las situaciones actuales, estimaciones de intervalos temporales, estudios a largo plazo de reanálisis). En la parte izquierda: requisitos de observación para un modelo de sistema terrestre integral, que además de los datos *in situ* incluye más datos satelitales actuales y estimados (satélites operacionales y de investigación). Las misiones operacionales y de investigación actuales cumplen muchos de los requerimientos espaciales, algunos de los cuales todavía no se cumplen. Las misiones satelitales se clasifican en función del ámbito geofísico observado (atmósfera, océano, superficie terrestre, tierra sólida) y por la frecuencia electromagnética utilizada por las misiones. En el centro: las herramientas que requieren un centro de pronóstico de sistemas terrestres (por ejemplo, ECMWF) y de sitios especializados de clientes/sitios asociados, con la finalidad de transformar las observaciones en información. Fuente: Hollingsworth *et al.*, 2005.

El uso racional del sistema de la Tierra, en lo que respecta tanto a los aspectos humanos como a los naturales, exige tener información relevante y contar con ella en el momento oportuno. A pesar de todos los esfuerzos realizados, la situación actual de las observaciones de la Tierra está muy lejos de ser satisfactoria, sobre todo, en lo que afecta a la coordinación y el intercambio de datos entre países, organizaciones y disciplinas a efectos de satisfacer las necesidades del desarrollo sostenible.

El GEOSS está construyendo y añadiendo valor a las observaciones sobre la Tierra planificadas y existentes conectándolas unas a otras; en ese marco aparecen las reflexiones sobre la interoperabilidad. En efecto, sistemas y componentes tienen que permitir que los datos y la información generados puedan combinarse. No obstante, las cuestiones que esto plantea y los mecanismos para ponerle remedio se analizarán en el punto relativo a la interoperabilidad.

Convertirse en un sistema de sistemas solo será una realidad si adopta los estándares apropiados para las *interfaces* a través de los distintos componentes de que dispone para el intercambio de datos e información. Está pensado para ser una pasarela entre los que generan datos geoespaciales y los usuarios finales de los mismos con la finalidad de potenciar la relevancia de las observaciones de la Tierra para cuestiones globales (Giuliane *et al.*, 2013; GEO Secretariat, 2005). Cada vez es más necesaria la integración de datos.

Los datos de observación de la Tierra se están multiplicando de una forma excepcionalmente rápida: los nuevos datos de satélites, los sensores de control remoto, etc. (Nativi *et al.*, 2015). Todo ello proporciona una ingente cantidad de información.

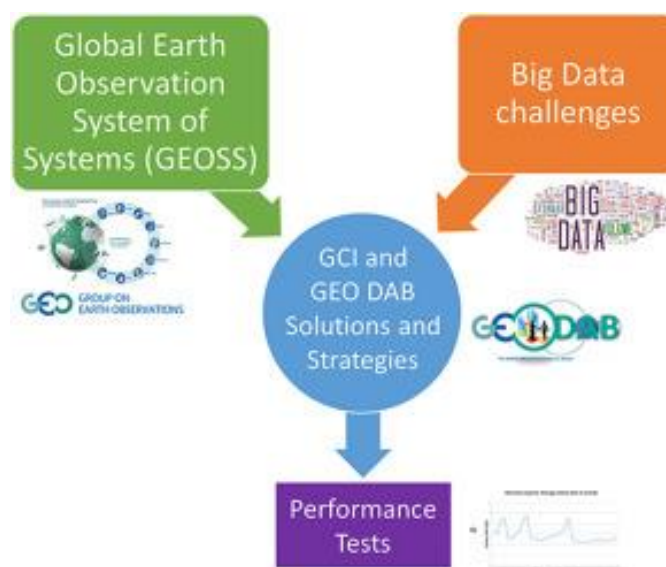


Figura 40. Representación gráfica. Fuente: Nativi, Mazzetti, Santoro, Papeschi y Craglia, 2015.

Es una magnífica oportunidad de promover el conocimiento de la Tierra, aunque, al mismo tiempo, plantea importantes retos tanto a científicos como a expertos en tecnología (Bargellini *et al.*, 2013).

4.1.2.1. El objetivo del GEOSS

La realización de un esfuerzo global, integral y sostenible como el GEOSS exigía resolver las siguientes cuestiones, afrontadas en el plan de implementación a diez años (GEO Secretariat, 2005):

1. Acuerdos para que los sistemas sean interoperables y para poder compartir datos

La capacidad de combinar datos de distintas fuentes incrementa sustancialmente el número y el tipo de observaciones disponibles para el análisis, así como su cobertura temporal y espacial, con un coste efectivo. El GEOSS proporcionaba un mecanismo a través del cual se podía negociar la forma de compartir datos de forma total o parcial. El GBIF, sobre el que se volverá en detalle en el punto de biodiversidad informática, es un claro ejemplo del poder que tiene este enfoque.

2. Optimizar las estrategias de observación

En cada tópico a estudiar, era necesario recoger un mínimo de datos. En caso de no existir colaboración, cada sistema de observación necesitaba realizar los cálculos de forma individual y desplegar su propia red y sistemas asociados para recabar los datos necesarios. La cooperación generaba sinergias y eliminaba duplicidades y los avances científicos eran mucho más rápidos usando sistemas híbridos de observación *in situ* y vía satélite.

Esta optimización solo podía conseguirse si existía un nivel de conocimiento de la distribución y de las dinámicas del sistema en observación, por eso, era clave mantener relaciones interactivas entre los sistemas de observación, los programas de investigación y las comunidades de usuarios.

3. Encontrar formas de cooperación para colmar lagunas

Muchos procesos de los sistemas de la Tierra operan a gran escala. Las deficiencias en la observación de un área tienen impacto en otras áreas. Las observaciones dentro de un país correspondían a ese país en concreto; sin embargo, en la Tierra existen muchas áreas fuera de la soberanía de un país, por ejemplo: el espacio, los océanos, o la Antártida. La observación de estas áreas reporta beneficios para todos y debe compartirse. El GEOSS proporciona mecanismos para identificar lagunas y movilizar recursos para cubrirlas.

4. Compromisos con la adecuación y continuidad de las observaciones

Ninguna de las acciones anteriores será efectiva a largo plazo salvo que exista un compromiso fundamental de continuar con las observaciones a un nivel aceptable de precisión y cobertura. Formar parte del GEOSS significa aceptar esta necesidad de adecuación y continuidad. En algunos casos, lograrlo exigirá la dedicación de recursos adicionales.

5. Transferencia de datos y diseminación

Hacer posible que todos los usuarios de forma global reciban los datos relevantes en el momento oportuno es imperativo para maximizar la explotación exitosa de las observaciones de datos y productos. Esto lleva aparejado recoger datos globales, particularmente de redes *in situ*, transferir los datos y productos entre agencias responsables para las observaciones y productos y diseminar datos y productos a los usuarios. En algunos ámbitos, se ha invertido mucho esfuerzo en la consecución de este objetivo.

La tecnología que hace esto posible está desarrollándose a mucha velocidad e incluye el uso de Internet, los satélites de telecomunicaciones y las conexiones de banda ancha. Existen muchos problemas de acceso a los datos en los países en vías de desarrollo, en particular en las áreas rurales, que los satélites de telecomunicaciones pueden resolver, en parte, aunque acarrearán problemas de coste.

6. Colaboración en la generación de capacidad

La forma más eficiente de mejorar la cobertura geográfica del sistema de observación de la Tierra es promover la amplia participación de todos los países, tanto en calidad de proveedores como en calidad de usuarios de dicha información. Algunos tipos de observaciones son difíciles de justificar, en particular en los países en vías de desarrollo, en términos de beneficio inmediato local. Por ende, son de baja prioridad en estos países. Los beneficios de hacer y compartir estas observaciones son colectivos y existen significativas ventajas en generar capacidad de forma colectiva. La capacidad se genera con e-infraestructuras, recursos humanos y capacidad institucional.

7. Armonización de métodos y aplicación de estándares de observación

La combinación de datos de distintas fuentes es esencial para progresar en nuestro conocimiento sobre el sistema de la Tierra, aunque este, en muchos casos, sufre limitaciones como consecuencia de observaciones incompatibles, inexistencia de estándares e insuficiencia de capacidad de síntesis de datos operacionales. El GEOSS proporciona un sistema para lograr la convergencia y la organización de los métodos de observación, del uso de estándares y de referencias y de la promoción de intercalado y síntesis de datos operacionales.

Una de las claves para que funcione es la interoperabilidad de los sistemas, datos, etc. (Onsrud *et al.*, 2010). Sobre el tema de la interoperabilidad se entrará con más detalle en el punto relacionado con la misma. El GEOSS está enfocado en nueve áreas denominadas de beneficios sociales que se recogen en el siguiente gráfico.



Figura 41. Imagen extraída de la página principal de GEOSS. Fuente: <http://www.earthobservations.org/geoss.php>.

Su finalidad y visión es «conseguir un futuro en el que las decisiones y acciones que beneficien a la humanidad se tomen de una forma coordinada, integral y sostenible de acuerdo con la información extraída de las observaciones de la tierra». Va un paso más allá, es una importante contribución para alcanzar los Objetivos de Desarrollo del Milenio.

Como se comentaba, al inicio, el GEOSS identifica nueve áreas de beneficios sociales, desde la gestión de catástrofes hasta la agricultura sostenible pasando por la variabilidad y el cambio climático que serán las prioritarias a la hora de recopilar la información. Siendo consistente con estos objetivos, también tiene la función de sensibilizar en relación con la necesidad de que existan más esfuerzos de observación de la Tierra y en promover un mejor uso de las decisiones adoptadas y los beneficios para la sociedad. Es de esperar que contribuya a evitar innecesarias duplicaciones de esfuerzos en la recopilación de los datos principales y de las lagunas de información. De la misma manera, la expectativa es que ayude a los Gobiernos y a las organizaciones que forman parte de la organización a planificar

nuevas inversiones a la hora de compartir los datos de observación de la Tierra y otros datos relacionados con estos.

El plan decenal de ejecución del GEOSS reconoció expresamente la importancia de compartir los datos para lograr su visión y conseguir los beneficios para la sociedad. Este punto se abordará en el apartado de análisis de los *data sharing principles* o principios de intercambio.

4.1.2.2. Los *data sharing principles*

La necesidad de integrar el conocimiento de las diversas disciplinas dedicadas al estudio de las partes constitutivas del complejo sistema de la Tierra es cada vez mayor. Con la aparición del GEOSS, que reúne miles de sistemas de observación de la Tierra previamente aislados, la necesidad de establecer métodos que ayuden en la integración de variados sistemas de información disciplinaria deviene, si cabe, aún más urgente. El Grupo sobre Observaciones Terrestres (GEO) fue creado para supervisar la creación del GEOSS para avanzar en la convergencia de los sistemas de observación de la Tierra basados en acuerdos de interoperabilidad decididos por consenso (Singh Khalsa *et al.*, 2009).

Existía un plan en el GEOSS donde se pusieron de manifiesto los siguientes principios que servirían de base para el intercambio de los datos:

1. Compartir de forma total y abierta los datos, metadatos y productos, reconociendo los instrumentos internacionales relevantes y las políticas nacionales así como la regulación.
2. Poner a disposición todos los datos compartidos, metadatos y productos, en los periodos de tiempo más cortos y con el mínimo coste.
3. Entregar todos los datos compartidos, metadatos y productos sin cargo o con un cargo que no supere el coste de reproducción para promover la investigación y la formación.

En 2013 un nuevo informe GEO señaló la necesidad de mejorar los marcos de interoperabilidad: la finalidad última era el acceso de los usuarios a la ingente cantidad de datos ahora existentes, recopilados, almacenados, etc. El trabajo aislado de los distintos sistemas que recogían datos medioambientales de forma aislada dificultaba, por lo tanto, encontrar, acceder o usar los contenidos de estos sistemas debido a la incompatibilidad de los sistemas y a las inconsistencias de formatos y de modelización de datos (Bernard *et al.*, 2005). Además, el intercambio de datos tendía a ser insuficiente fruto de las distintas políticas de intercambio de datos (Giuliane *et al.*, 2013).

Una de las mayores limitaciones a la hora de compartir ampliamente los datos de observación de la tierra es la inexistencia de un sistema web transparente y consistente que permita a los usuarios acceder legalmente y usar los datos de observación de la Tierra o cualesquiera otros sin tener el permiso de los autores o investigadores de las condiciones de uso caso por caso (Onsrud *et al.*, 2010).

Existe otro sistema concebido como un posibilitador de la nueva era de observación de la Tierra (Moodley *et al.*, 2008). El sensor web como sistema integra conjuntos de datos, fuentes, sensores, modelos de simulación y procesadores de datos heterogéneos y globalmente instalados. Sirve de piedra angular para el GEOSS. Los beneficios de este incluyen el uso de recursos mediante la reconfiguración de los sensores elegidos para dar soporte a las nuevas cuestiones que plantea la ciencia.

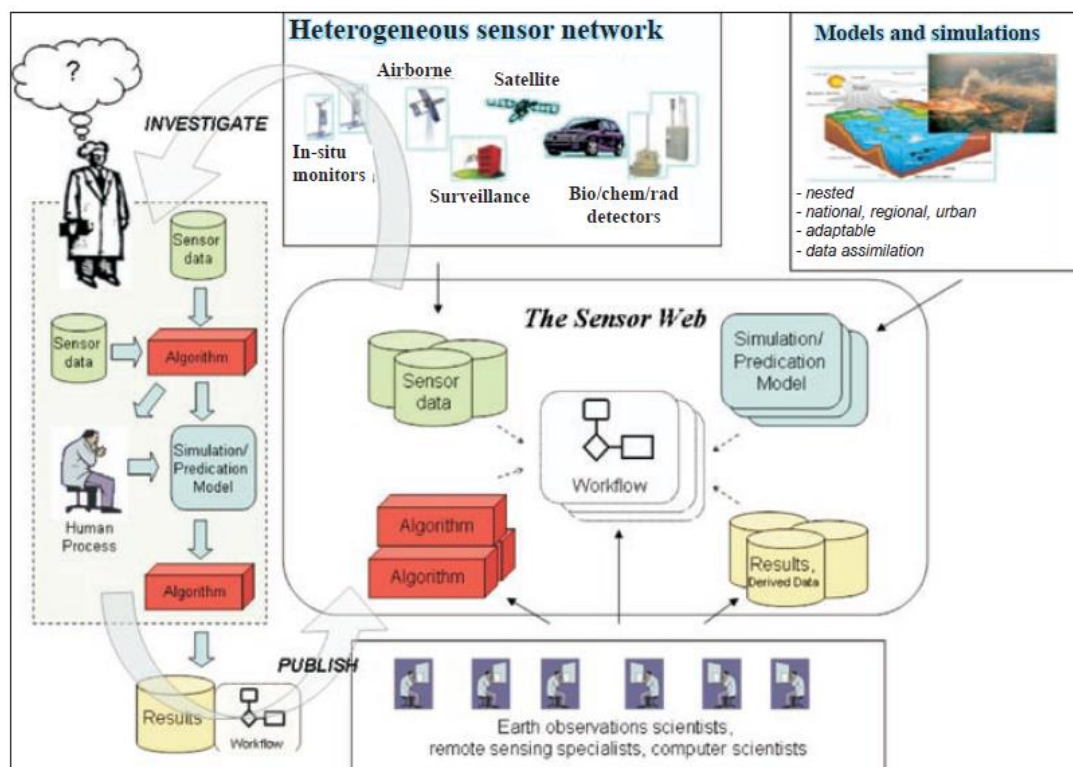


Figura 42. Arquitectura conceptual de alto nivel para la plataforma SW4SW. Fuente: Moodley *et al.*, 2008.

El intercambio de datos, por lo general, es el prerequisite para desarrollar el acceso a los datos de biodiversidad, aunque lograrlo no resulte tan fácil: que exista intercambio no es suficiente para que los datos de biodiversidad sean accesibles. Existen una multiplicidad de razones: falta de estándares adecuados, incentivos y recursos, y también inexistencia de políticas de intercambio de datos a nivel nacional y europeo (Wetzel *et al.*, 2015).

Por eso, el hecho de que el GEO en 2010 aceptara un conjunto de principios de alto nivel de intercambio de datos como punto de partida del GEOSS es un logro importante, a pesar de que la implementación esté resultando todo un reto.

En 2010, el GEO creó un Data Collection of Open Resources for Everyone (Data-CORE) en el ámbito del GEOSS aplicando sus principios de intercambio de datos (GEO Data Sharing Working Group, 2014). La finalidad del GEOSS Data-CORE era promover el acceso a los conjuntos de datos relacionados con la observación de la Tierra y hacer posible el uso y la reutilización de los datos sin restricciones respecto a lo que afecta a este uso y reutilización.

En la asamblea celebrada en 2011, y la posterior actualización de 2012, establece las opciones legales en relación con el intercambio de datos recogiendo en un *White Paper* que analizará a continuación. En base a los estudios realizados para este documento, define la interoperabilidad legal. Considera que la interoperabilidad legal entre una multiplicidad de conjuntos de datos procedentes de distintas fuentes se da cuando:

- Las condiciones de uso sean clara y rápidamente definibles en relación con cada uno de los conjuntos de datos.
- Las condiciones de uso legal impuestas a cada uno de los conjuntos de datos permitan la creación y el uso de productos combinados o derivados.
- Los usuarios tengan la posibilidad de acceder de forma legal y usar cada conjunto de datos sin necesidad de obtener la autorización de los creadores de los mismos caso por caso, suponiendo que cada uno de los conjuntos de datos cumpla las condiciones acumuladas de uso.

Por consiguiente, el estatus más adecuado para que se cumplieran las condiciones de interoperabilidad era que los datos se pusieran en el dominio público. Es, sin duda, la forma más eficaz de permitir la reutilización, disseminación y la interoperabilidad legal. Esto es lo que mismo que decir que en el marco del GEOSS se establece una política *open data by default*.

Los derechos derivados del *copyright*, de la propiedad intelectual o industrial o, más concretamente, la protección *sui generis* de las bases de datos nacen de forma automática y permanecen vigentes hasta el momento en que expira el término de protección o bien hasta el momento en que se renuncie a estos derechos. Por ello, es necesario que existan mecanismos jurídicos, regulatorios, políticas o mecanismos de subvención con la finalidad de poder eliminar la protección o que se pueda renunciar a ella y permitir el uso sin restricciones de los datos, su disseminación y su

reutilización. Puede ocurrir también que determinadas organizaciones privadas renuncien a estos derechos siempre y cuando la ley que los regula lo permita.

Lo ideal es que los datos que se pongan en el dominio público incluyan un aviso en sus metadatos o en la base de datos del propietario que recoja este extremo con la finalidad de que sus potenciales usuarios puedan conocerlo. Este aviso promoverá el uso de estos datos y la interoperabilidad legal dentro del marco del GEOSS Data-CORE.

Algunas herramientas contribuyen a estos fines, como es el caso de las licencias Creative Commons (CC). De hecho, algunos autores se plantean las licencias como una alternativa al *copyright* (Xalabarder Plantada, 2006).

Una forma de contribuir a la interoperabilidad legal de los datos y, por lo tanto, de dar soporte a las áreas de interés del GEOSS es que el Pleno del GEO apruebe el uso de licencias o renunciaciones (*waivers*) estandarizadas y aceptadas o de cualesquiera otras licencias parametrizadas que reúnan todas las condiciones de acceso y de uso sin restricciones del GEOSS Data-CORE.

Las licencias de uso común y los *waivers* también ayudarían a promover la contribución de las bases de datos a través del GEOSS Data-CORE, porque en la mayoría de las jurisdicciones las compilaciones de datos relevantes para el GEOSS no tienen la condición de dominio público. Este paso también será útil para los miembros y las organizaciones participantes que estén dispuestos a compartir datos como parte del GEOSS Data-CORE, ya que economizará los recursos que necesitarían gastar en el desarrollo de tales licencias por sí mismos.

El Grupo de Trabajo de Intercambio de Datos del GEO recomienda usar solo instrumentos estándar para ayudar a asegurar la interoperabilidad legal de los datos. Si no se pueden utilizar renunciaciones estándar o licencias de uso común; el proveedor de datos puede considerar la adopción de una licencia de uso personalizado o de una licencia de uso común. Dicha renuncia o licencia debe ser compatible con los principios del GEOSS *Data-CORE* sin restricciones de reutilización, permitiendo el registro de usuarios, las condiciones de atribución y los cargos de recuperación de costes marginales. Además, debe:

1. Ser válida según cuantas más leyes de jurisdicciones sea posible.
2. Ser clara y comprensible para el proveedor o usuario de datos.
3. Ser fácil de encontrar y reconocer.
4. Permitir incorporarse en los datos como metadatos legibles máquina-máquina siempre que sea posible.

5. Estar disponible en diferentes idiomas, como mínimo en la lengua o los idiomas del país u organización que pone a disposición los datos, así como en inglés.
6. Mantenerse bajo el control legal de los proveedores de datos, y no del GEO o del GEOSS.

Podrán, además, usarse cualesquiera otros mecanismos que garanticen el intercambio y uso abierto de los datos. Después de 2015 (Data Sharing Working Group, DSWG, of the Group on Earth Observations, GEO, 2014) los principios de intercambio de datos del GEOSS (Group on Earth Observation, s. f.) quedan recogido de la forma que se expone a continuación.

El GEO reconoció que los beneficios sociales derivados de las observaciones de la Tierra solo podrían lograrse plenamente mediante el intercambio de datos, información, conocimiento, productos y servicios. Por lo tanto, el GEO promovió principios fundamentales para este intercambio, ampliando la tendencia hacia la apertura de datos en todo el mundo. Así, a medida que se embarcaron en su segunda década, el GEO ahora tiene como objetivo implementar los siguientes principios de intercambio de datos del GEOSS:

- Los datos, los metadatos y los productos se compartirán como datos abiertos de forma predeterminada, haciéndolos disponibles como parte de la recolección de datos del GEOSS de recursos abiertos para todos (Data-CORE) sin cargo ni restricciones de reutilización, con sujeción a las condiciones de registro y atribución en la reutilización de los datos.
- Cuando los instrumentos internacionales, las políticas nacionales o la legislación impidan el intercambio de datos como abiertos, estos deberán estar disponibles con las mínimas restricciones de uso y con un coste no superior al de reproducción y distribución.
- Todos los datos, productos y metadatos compartidos estarán disponibles con el mínimo retraso posible o, en positivo, lo más rápidamente posible.

La base sobre la que se sustentan estos principios es la siguiente:

- Afirmar que el intercambio de datos como parte del GEOSS Data-CORE es el estándar por defecto para el GEO eleva el estado de este mecanismo, así como su importancia global para el éxito de la operación del GEOSS y los logros de los objetivos del GEO, observaciones que se destacan en el documento visión para el GEO 2025 adoptado por la Sesión Plenaria del GEO X (Anon., 2014).

- La referencia al término «datos abiertos» proporciona un contexto para la interpretación de las condiciones de uso pertinentes a los datos compartidos como parte del GEOSS Data-CORE, así como a los principios de intercambio de datos del GEOSS en línea con los desarrollos internacionales, regionales, etc.
- La opción de compartir datos a través del GEOSS con restricciones de uso se presenta como una desviación del mecanismo por defecto, con el énfasis en imponer tan pocas restricciones en el uso de datos compartidos como sea posible. Este cambio en el énfasis reconoce mejor las motivaciones para GEOSS: fomentar y facilitar la reutilización de datos y productos del GEO, así como ayudar a tomar decisiones informadas dentro de nueve áreas de beneficio social.
- La definición de *open data* conlleva que los datos se comparten gratuitamente, con cualquier finalidad y con cualquier usuario. Esto refleja la tendencia actual de muchos Gobiernos hacia los datos abiertos y es conforme a los objetivos del GEO de fomentar el intercambio de datos con el fin de abordar los objetivos sociales establecidos y promover los beneficios económicos. La redacción actual del tercer principio que limita el reparto gratuito de los gastos de investigación y educación es menos apta para alcanzar estos objetivos.
- En el libro blanco *Mecanismos para compartir datos como parte del GEOSS Data-CORE*, aprobado por el Pleno del GEO en noviembre de 2014, se presentan y analizan diversos mecanismos legales para la puesta a disposición de datos como parte del GEOSS Data-CORE.

El informe sobre el valor del intercambio abierto de datos fue preparado por primera vez para la Sesión Plenaria GEO-XII por la organización participante del GEO CODATA (Comité de Datos sobre Ciencia y Tecnología del ICSU) (GEO XIII, 2014).

A través de la explicación de los diversos beneficios de los datos de observaciones abiertas de la Tierra, el informe está diseñado para facilitar el proceso de transición de políticas de datos restringidos a políticas más abiertas para los datos del Gobierno. En la Sesión Plenaria GEO-XII, el informe recibió retroalimentación positiva de los países miembros del GEO y de las organizaciones participantes, quienes también expresaron su voluntad de aportar estudios de caso adicionales. Por lo tanto, se decidió mantener este informe como un documento vivo para el registro como versión 1. La comunidad del GEO proporcionará actualizaciones periódicas, ejemplos y estudios de casos.

El GEOSS es un ejemplo de los esfuerzos que se están realizando a nivel global a la hora de adoptar tecnologías estándar y especificaciones en lo que respecta a los retos de interoperabilidad (Nativi *et al.*, 2011). Como puede verse, el énfasis en el *open data* se está acentuando en estos últimos años.

4.1.2.3. Los *data management principles*

El GEOSS da un paso más en lo que respecta a los principios de intercambio de datos y propone la elaboración de un conjunto de principios que no afecten solamente al intercambio de los datos, sino también a la gestión de los mismos (GEO Data Management Principles Task Force, 2015). Estos tuvieron su origen en el documento elaborado en 2014 en relación con el intercambio de datos (GEO-XI, 2014).

Los *Data Management Principles* del GEOSS se basan en los *Data Sharing Principles* del GEOSS y esbozan lo que se requiere en términos de gestión de datos para permitir que los datos se compartan como *open data*, con rapidez y con un coste mínimo. Una buena gestión de datos implica una serie de actividades que aseguran que los datos sean accesibles y puedan ser entendidos y utilizados. Además deberá asegurarse su cuidado a largo plazo (G8 Science Minister's, 2013).

Para el GEO es prioritario fomentar que las organizaciones integrantes del GEOSS implementen los principios. Además, señalaron la necesidad de establecer directrices para su implementación.

Respecto a la capacidad de ser encontrados (*discoverability*):

Data Management Principle 1. Los datos y todos los metadatos asociados deben poder encontrarse mediante catálogos y motores de búsqueda. Indicarán claramente las condiciones de acceso y de uso de los datos, incluidas las licencias.

Respecto a la accesibilidad (*accessibility*):

Data Management Principle 2. Los datos serán accesibles mediante servicios en línea, incluyendo, como mínimo, la descarga directa, pero preferiblemente mediante servicios personalizables para la visualización y computación.

Respecto a la usabilidad (*usability*):

Data Management Principle 3. Los datos se estructurarán utilizando codificaciones ampliamente aceptadas en la comunidad de usuarios a quien se dirige y alineadas con las necesidades de la organización y los métodos de observación; se dará preferencia a los estándares internacionales abiertos o no registrados.

Data Management Principle 4. Los datos se documentarán de forma exhaustiva, incluyendo todos los elementos necesarios para acceder, utilizar, comprender y procesar, preferiblemente mediante metadatos estructurados formales basados en estándares internacionales o aprobados por la comunidad.

Data Management Principle 5. Los datos incluirán metadatos de procedencia que indicarán el origen y el historial de procesamiento de observaciones primarias y productos derivados, para asegurar la trazabilidad completa de la cadena de productos.

Data Management Principle 6. La calidad de los datos deberá controlarse y los resultados del control de calidad deberán aparecer señalados en los metadatos. Los datos que estén disponibles antes de haberse realizado el control de calidad aparecerán marcados en los metadatos señalando este extremo.

Respecto a la preservación (*preservation*):

Data Management Principle 7. Se protegerán los datos de su pérdida y se conservarán para usos futuros. La planificación de la conservación deberá hacerse a largo plazo e se incluirán directrices para prevenir las pérdidas, calendarios de retención y procedimientos de eliminación o transferencia.

Data Management Principle 8. Los datos y los metadatos asociados contenidos en los sistemas de gestión de datos se verificarán periódicamente para garantizar la integridad, la autenticidad y la legibilidad.

Respecto a la conservación (*curation*):

Data Management Principle 9. Los datos se gestionarán para realizar correcciones y actualizaciones de conformidad con las revisiones, y para permitir el reprocesamiento, según proceda. En los casos en los que sea adecuado, seguirán los procedimientos establecidos y acordados.

Data Management Principle 10. A los datos se les asignarán identificadores persistentes apropiados para permitir que los documentos citen los datos en los que se basan y para permitir que los proveedores de datos reciban el reconocimiento del uso de sus datos.

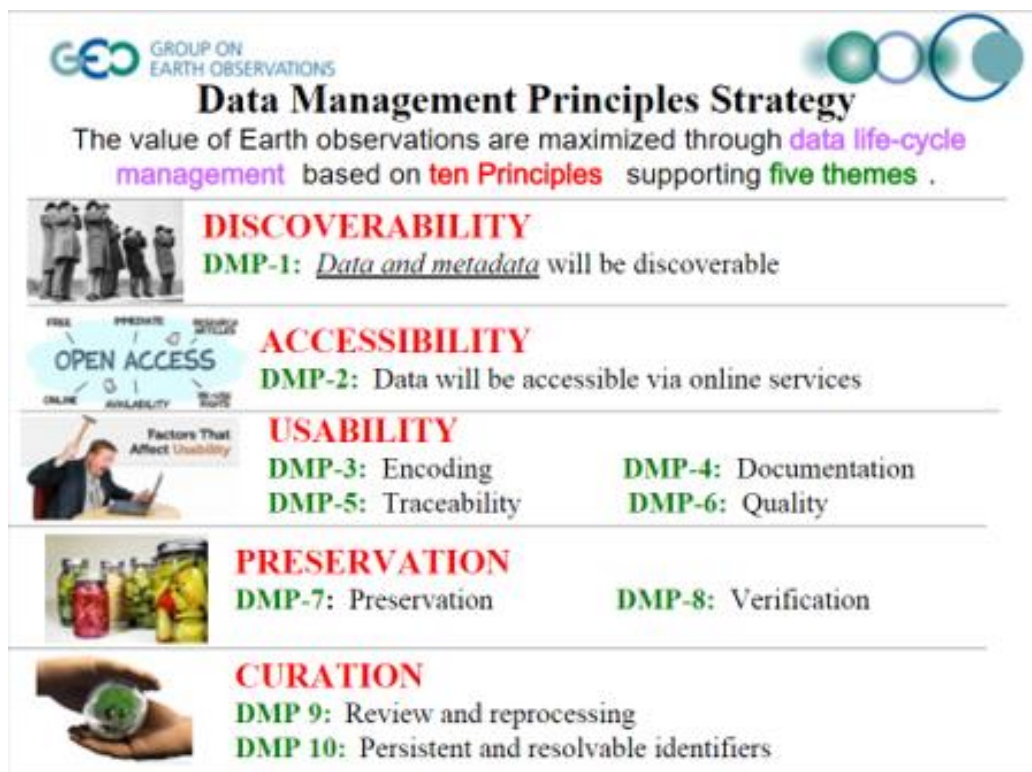


Figura 43. Estrategia de *Data Management Principles*. Fuente: Alonso García, 2016a.

No se han hecho todavía grandes desarrollos científicos o análisis sobre estos principios de gestión. Lo que sí es evidente es que se van consolidando las políticas de *open data*, *open science* y directrices de gestión con pautas de uso de los mismos.

4.2. Committee on Data for Science and Technology (CODATA). *Policies for open access to research data*

El Committee on Data for Science and Technology (Comité de Información para Ciencia y Tecnología, en adelante, CODATA por sus siglas en inglés) es un comité interdisciplinar del Scientific Committee of the International Council for Science (ICSU, por sus siglas en inglés y traducido en español como Consejo Internacional para la Ciencia) fue creado hace más de cuarenta años (Committee on Data of the International Council for Science, s. f.).

CODATA trabaja para mejorar la calidad, fiabilidad, gestión y accesibilidad de los datos relevantes en todos los ámbitos de la ciencia y de la tecnología. Es un recurso que proporciona a científicos e ingenieros acceso a actividades de datos internacionales para aumentar la concienciación, la cooperación directa y el desarrollo de nuevo conocimiento. Fue creado en 1966 por el ICSU para promover y fomentar, a nivel mundial, la recopilación, evaluación y difusión de datos numéricos confiables de importancia para la ciencia y la tecnología.

Se ocupa de todos los tipos de datos resultantes de mediciones experimentales, observaciones y cálculos en todos los campos de la ciencia y la tecnología, incluyendo las ciencias físicas, la biología, la geología, la astronomía, la ingeniería, la ciencia ambiental y la ecología, entre otras disciplinas. Hace especial hincapié en los problemas de gestión de datos comunes a las distintas disciplinas y en los datos utilizados fuera del ámbito en el que se generaron.

Los objetivos de CODATA son:

- Mejorar la calidad y la accesibilidad de los datos, así como los métodos por los cuales se obtienen, gestionan, analizan y evalúan, con un énfasis particular en los países de desarrollo.
- Facilitar la cooperación internacional entre aquellos que recaban, organizan y utilizan datos.
- Promover una concienciación cada vez mayor dentro de la comunidad científica y técnica sobre la importancia de dichas actividades.
- La consideración del acceso a datos y de las cuestiones de propiedad intelectual.

En resumen, la razón de ser de CODATA es contribuir al desarrollo y progreso de la ciencia y de la tecnología mediante el desarrollo y el intercambio del conocimiento en relación con los datos y con las actividades relacionadas con estos.

Para alcanzar estos objetivos CODATA recurre a distintos mecanismos como pueden ser: *tasks groups*, *working groups*, conferencias, actividades en los Estados nacionales, talleres, publicaciones, cooperación con otras organizaciones con intereses comunes.

El logro de esos objetivos es un fin en desarrollo continuo. Los métodos utilizados para alcanzar dichos objetivos están mejorando y evolucionando continuamente. Es importante enfatizar que las actividades relacionadas con los datos no se limitan a una disciplina científica específica, sino que se relacionan con actividades de datos en todas las áreas de la ciencia y la tecnología, desde la biología hasta el cambio global, desde las ciencias físicas hasta la ingeniería. CODATA abarca diferentes disciplinas mediante su creciente número de miembros dentro de la familia, donde cada uno resalta sus diferentes necesidades en lo que afecta a los datos, comparte conocimiento sobre sus respectivas actividades de datos e identifica áreas comunes de interés; estas últimas son cuatro y todas ellas sirven de soporte a su finalidad fundamental, que es la de potenciar la cooperación internacional en relación con los datos científicos y técnicos:

- Patrocinar la Conferencia Internacional Bianual en datos de CODATA, que reúne a trescientos expertos en datos de todo el mundo.
- Celebrar reuniones especializadas de científicos expertos en datos, para resolver alguna cuestión específica de alguna disciplina o tema.
- Publicar estudios en la gestión de datos, compilaciones de datos, encuestas de actividades de datos y actas de conferencias.
- Patrocinar los *tasks groups*, *working groups*, comisiones y otros grupos que analicen cuestiones concretas relacionadas con los datos como puede ser:
 - La coordinación de proyectos de datos multinacionales.
 - El establecimiento de formatos estándar para promover el intercambio de datos y la compatibilidad.
 - Directrices de presentación de los datos en la literatura primaria y en los bancos de datos.
 - Suministrar información de recursos donde puedan encontrarse datos confiables.
 - Educación y formaciones.
 - Preparación de conjuntos de datos clave donde es deseable que exista un uso internacional consistente.
 - Organización de conferencias y de talleres.

Los comités nacionales de CODATA organizan, con frecuencia, actividades relacionadas con los datos en el ámbito nacional. Sin embargo, gran parte del trabajo más importante de CODATA está fuera de su actividad normal por el mero hecho de proporcionar un entorno en el que pueden interactuar expertos en datos de distintos países, cooperar de forma directa, desarrollar colaboraciones bilaterales fuera de CODATA e intercambiar tanto ideas como conocimiento.

Una de las principales fuerzas motrices para la creación de CODATA fue el hecho de que la gestión de los datos científicos ya no podía dejarse a los esfuerzos especiales de algunas personas dedicadas. La rápida expansión de la investigación científica que había ocurrido continuamente desde el final de la Segunda Guerra Mundial trajo consigo una expansión correspondiente en la cantidad de datos generados. Si se quiere aprovechar plenamente esta inversión en investigación, se necesitan mejores mecanismos para evaluar, almacenar, recuperar y difundir los datos. Si bien las organizaciones nacionales llevan a cabo algunas de estas funciones, CODATA es la organización coordinadora internacional que anima, coordina e intenta evitar la duplicación de tales esfuerzos.

CODATA está abierta a nuevos proyectos, especialmente en este momento de creciente importancia de los datos debido al modelado informático y a Internet. Hoy en día, son cada vez más fuentes de datos de calidad y de origen desconocidos y se están haciendo fácilmente disponibles, a través de la World Wide Web. Proporciona, por lo tanto, un lugar en el que los expertos en datos internacionales que necesitan abordar la calidad de los datos y los problemas de acceso a los datos para convertir la revolución de la información en una fuerza positiva para el futuro. Los datos de hoy se convierten en los productos y procesos de mañana. Está preparado para ayudar a la ciencia y la tecnología a lograr un mañana mejor a través de mejores datos hoy.

Por lo tanto, nace con la misión de «reforzar la ciencia internacional en beneficio de la sociedad promoviendo una mejor *gestión y uso* de los *datos científicos y técnicos*» (Hodson, 2015).

Se creó un Comité de Data Policy, fundamentalmente, con la idea de ayudar a CODATA a alcanzar sus objetivos y de liderar en materia de políticas de datos a nivel internacional. Una de las funciones fundamentales que realiza es fomentar el desarrollo y la implementación de políticas de datos para construir la *open science*.

Ciertas observaciones en relación con las políticas de disponibilidad de datos son relevantes (Hodson, 2015):

- No deben existir interferencias burocráticas y deben contener sólidos principios de buena ciencia.
- Deben perseguir asegurarse de que los investigadores se benefician de hacer lo mejor por ejemplo en lo que respecta a las citas de datos.
- Los datos son mucho más efectivos cuando vienen y están refrendados por comunidades científicas, en este ámbito se trabaja así.
- Deben, de igual manera, proporcionar directrices de implementación.
- Necesariamente, deberán tener una dimensión internacional: la investigación es la tarea internacional colaborativa/competitiva por excelencia y esto hace que sea importante definir políticas marco.

Recoge, además, una tipología de políticas de datos que se muestra en el siguiente gráfico.

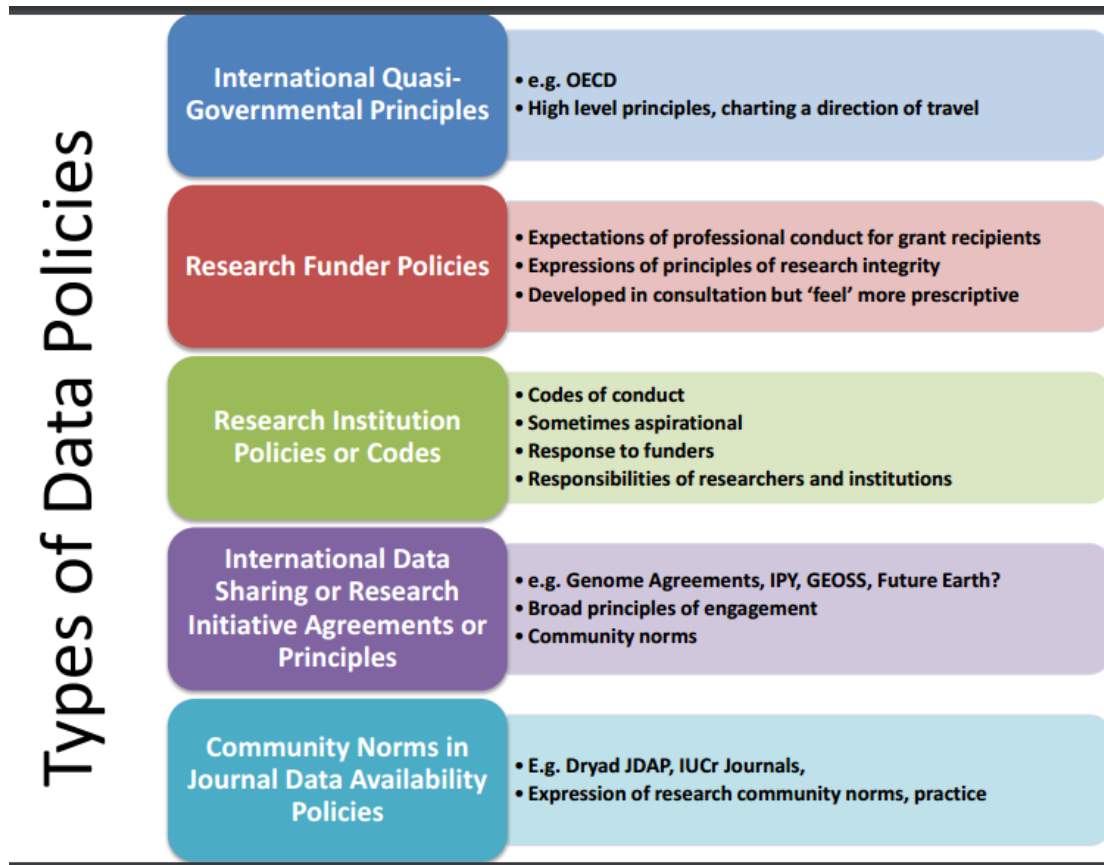


Figura 44. Representación gráfica. Fuente: Hodson, 2015.

Sin lugar a dudas, existe un nuevo paradigma: compartir los datos en abierto. A distintas velocidades se está implantando en las distintas ciencias y el resultado final será permitir a la ciencia el mayor avance ocurrido hasta ahora. La tendencia, aunque pueda parecer que avanza más despacio de lo esperable, es imparable.

4.3. La interoperabilidad legal

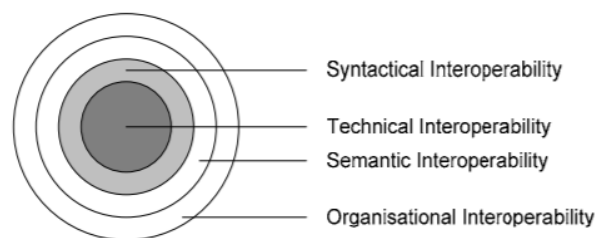


Figura 45. Distintos niveles de interoperabilidad. Fuente: Van der Veer, 2008.

La interoperabilidad tiene distintos niveles, previos a la construcción del concepto de interoperabilidad legal. Es clave, por lo tanto, abordar el significado de estos distintos niveles. Así, la interoperabilidad técnica de datos es la «propiedad de un producto o sistema [...] de trabajar con otros productos o sistemas, actuales o

futuros sin ninguna restricción en su acceso o implementación» (Alonso García, 2016*b*). Por consiguiente, la interoperabilidad técnica es la base y presupuesto de todas las demás.

El siguiente nivel de interoperabilidad es la sintáctica que ocurre cuando dos o más sistemas se pueden comunicar entre sí. Se logra con formatos de datos, protocolos y estándares específicos XML o SQL u otros de menor nivel (Alonso García, 2016*b*). En un sistema de base web, la interoperabilidad sintáctica supone estandarizar la comunicación entre un *software* de cliente y un servidor (Schaeffer *et al.*, 2012).

La Research Data Alliance (2015) define la interoperabilidad semántica como la capacidad de los servicios y de los sistemas de intercambiar datos de una forma útil y con sentido (Sudmanns *et al.*, 2018). Es un concepto mucho más complejo de lo que parece a primera vista, puesto que requiere que se entienda el sentido de los datos y de la información en el proceso de comunicación (Harvey *et al.*, 1999). Con la finalidad de facilitar que exista la interoperabilidad semántica, los *softwares* clientes que solicitan el proceso y los servidores, que llevan a cabo el proceso, acuerdan de forma clara y unívoca el contexto y los conceptos de las entidades y sus relaciones (Sudmanns, 2018).

Así, la interoperabilidad semántica se define como la capacidad que poseen los sistemas de información para intercambiar datos con significado unívoco y compartido. Es un requisito que permite a dichos sistemas ejecutar mecanismos de inferencia lógica, descubrimiento de nuevo conocimiento y federación de datos, con lo cual, puede decirse que pasan a denominarse «sistemas de información y del conocimiento». Por lo tanto, la interoperabilidad semántica no se refiere tan solo a la «sintaxis» y simple intercambio de los datos, sino también a la transmisión simultánea junto a estos de su significado (de ahí lo de «semántica»). Esto se lleva a cabo mediante la inclusión de los así llamados «metadatos» («datos de los datos», descriptores de estos en sí), enlazando cada dato a un determinado «vocabulario compartido» compuesto por dichos metadatos. Así pues, el significado de los datos es transmitido junto los datos en sí conformando un «paquete de información» que pueda ser interpretado de manera independiente por cualquier sistema de información y del conocimiento. Es este «vocabulario compartido», y sus enlaces asociados a una determinada «ontología» los que confieren las bases y la capacidad a las máquinas que integran dichos sistemas de información y del conocimiento de poder interpretar e inferir lógicamente a partir del significado de dichos datos. Una ontología es una especificación de una conceptualización, esto es, un marco común o una estructura conceptual sistematizada y de consenso no solo para almacenar la información, sino también para poder buscarla y recuperarla, de manera que una ontología define los términos y las relaciones básicas para la comprensión de un área del conocimiento, así como las reglas para poder combinar los términos para definir

las extensiones de este tipo de vocabulario controlado anteriormente referido (González Aranda *et al.*, 2009). Estos autores se basan en gran parte también en la tesis doctoral de Juan Miguel González-Aranda, *Creation and nurturing of Science & Technology cooperation tools by means of Virtual Communities of Practice by using Organizational Knowledge Management Systems and E-Research Collaboration Tools* (2012), a su vez resultado de la colaboración entre el CSIC (Rafael Rodríguez-Clemente) y el Departamento de Organización Industrial y Gestión de Empresas de la Escuela de Ingeniería de la Universidad de Sevilla (Sebastián Lozano-Segura) (Alonso García, 2016b).

La interoperabilidad, a fin de cuentas, es facilitar la «puesta en común» de los datos con la finalidad de que estos puedan reutilizarse. Sin embargo, la interoperabilidad es más que fontanería digital, es decir, que consiste asegurarse que bits y más bits fluyan adecuadamente. El interés en la interoperabilidad tiene su origen en el interés por compartir datos, por hacerlos reutilizables (Landsbergen y Wolken, 1998).

Los regímenes de derechos de propiedad intelectual o *copyright* parecen estar frenando, poniendo obstáculos al intercambio de datos a lo que se conoce con el nombre de interoperabilidad legal,⁴⁴ un concepto novedoso. En efecto, una condición necesaria para que pueda producirse la interacción de los datos es la interoperabilidad, es decir, permitir que una máquina pueda utilizar los datos que están en otra máquina.

La interoperabilidad legal implica que la búsqueda y el rastreo de las condiciones legales para los conjuntos de datos se realizará en línea a través de métodos automatizados. Cuando se combinen datos de fuentes múltiples, el conjunto de datos resultante deberá incorporar las restricciones acumuladas impuestas por cada una de las fuentes. Cuanto menor sea la cantidad de restricciones contenidas en los conjuntos de origen, menor será la cantidad de restricciones en los conjuntos de datos derivados. Los casos más sencillos de seguimiento e interoperabilidad legal ocurren cuando los conjuntos de datos se identifican afirmativamente como sin restricciones legales (Doldirina, Eisenstadt, Onsrud y Uhler, en prensa).

No obstante, la interoperabilidad no solo tiene una cara, pues compartir datos, ponerlos en común, tiene distintos ámbitos de aplicación que Landsbergen y Wolken (1998) dividen en cuatro: el político —lo que afecta a la privacidad—, el organizacional —falta de confianza—, el económico —dificultad en identificar los supuestos en que los bienes son un bien público— y el técnico —la incompatibilidad de *software* y *hardware* y la inconsistencia de las estructuras de datos—.

⁴⁴ En realidad, la interoperabilidad legal es una traducción literal del inglés, en español debería decirse «jurídica».

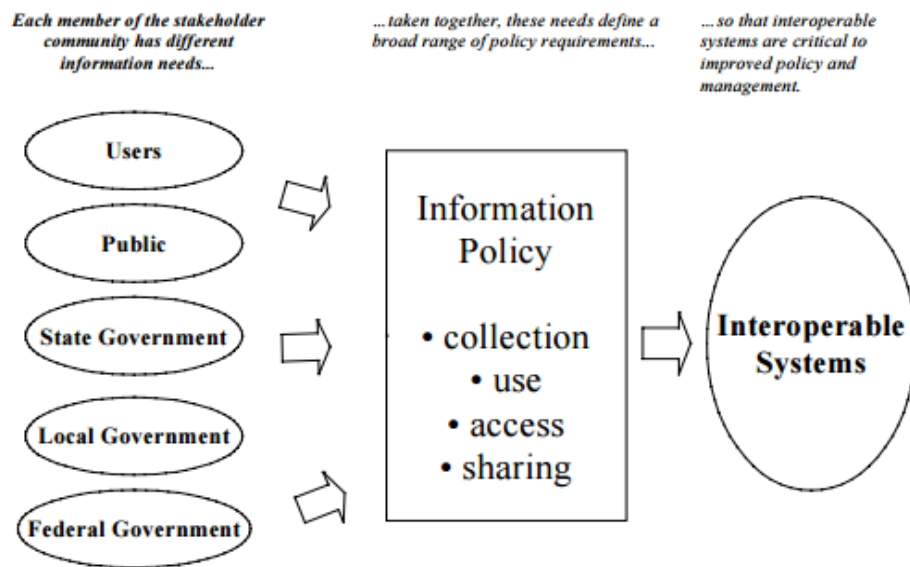


Figura 46. Representación gráfica. Fuente: Landsbergen y Wolken, 1998.

No son los únicos autores que dicen que la interoperabilidad tiene también que ver con cuestiones técnicas (Alonso García, 2016),⁴⁵ semánticas (Alonso García, 2016)⁴⁶

⁴⁵ Nota 10. La interoperabilidad técnica de datos es la «propiedad de un producto o sistema [...] de trabajar con otros productos o sistemas, actuales o futuros sin ninguna restricción en su acceso o implementación». Recuperado de <http://interoperability-definition.info/en>. (Consultado por última vez el 24 de abril de 2016).

⁴⁶ Nota 12. La interoperabilidad semántica se define como la capacidad que poseen los sistemas de información para intercambiar datos con significado unívoco y compartido. Es un requisito que permite a dichos sistemas ejecutar mecanismos de inferencia lógica, descubrimiento de nuevo conocimiento y federación de datos, con lo cual, puede decirse que pasan a denominarse como «sistemas de información y del conocimiento». Por lo tanto, la interoperabilidad semántica no se refiere tan solo a la «sintaxis» y simple intercambio de los datos, sino también a la transmisión simultánea junto a estos de su significado (de ahí lo de «semántica»). Esto se lleva a cabo mediante la inclusión de los así llamados «metadatos» («datos de los datos», descriptores de estos en sí), enlazando cada dato a un determinado «vocabulario compartido» compuesto por dichos metadatos. Así pues, el significado de los datos es transmitido junto los datos en sí conformando un «paquete de información» que pueda ser interpretado de manera independiente por cualquier sistema de información y del conocimiento. Es este «vocabulario compartido», y sus enlaces asociados a una determinada «ontología» los que confieren las bases y la capacidad a las máquinas que integran dichos sistemas de información y del conocimiento de poder interpretar e inferir lógicamente a partir del significado de dichos datos. Una ontología es una especificación de una conceptualización, esto es, un marco común o una estructura conceptual sistematizada y de consenso no solo para almacenar la información, sino también para poder buscarla y recuperarla, de manera que una ontología define los términos y las relaciones básicas para la comprensión de un área del conocimiento, así como las reglas para poder combinar los términos para definir las extensiones de este tipo de vocabulario controlado anteriormente referido (González Aranda *et al.*, 2009). Este autor se basa en gran parte también en la tesis doctoral del primero, Juan Miguel González-Aranda, *Creation and nurturing of Science & Technology cooperation tools by means of Virtual Communities of Practice by using Organizational Knowledge Management Systems and E-Research Collaboration Tools* (2012), a su vez resultado de la colaboración entre el CSIC (Rafael Rodríguez-Clemente) y el Departamento de

y sintácticas (Alonso García, 2016).⁴⁷ Por lo tanto, el punto de partida de cualquier interoperabilidad es que físicamente exista interoperabilidad técnica, semántica y organizativa y de ahí derivan los problemas que puedan plantearse. Su existencia exige posibilidad de interconexión entre los datos desde el punto de vista físico. Por consiguiente, para que pueda darse, tienen que cumplirse unas determinadas condiciones.

Aunque no podía dejar de hacerse referencia a los diferentes tipos de interoperabilidad, no son estas el objeto de esta tesis. Por lo que se analizarán en detalle los problemas que plantea la interoperabilidad legal.

Existe un concepto *narrow* de interoperabilidad legal que consiste en que hay que tener suficientes metadatos como para que las propias máquinas identifiquen si pueden reutilizarse esos datos o existen restricciones que lo impidan. Además, la propia máquina tiene que avisar al usuario de cómo obtener los derechos de reutilización. De esta manera, se pueden eliminar las barreras legales para dicha reutilización.

El hecho de que se haya dado importancia a la interoperabilidad legal es el resultado de toda la evolución que ha sucedido tanto en Europa como en Estados Unidos. La era del conocimiento no permite dejar esto a un lado, pues la proliferación del conocimiento en los distintos ámbitos está cambiando de enfoque.

Una parte importante del problema reside en que las distintas organizaciones funcionan dentro de un marco territorial cerrado y nacional. Las primeras dudas surgieron cuando las organizaciones empiezan a ser supranacionales, cuando el derecho internacional todavía tenía instrumentos para regularlas. Actualmente, el derecho internacional ya no es suficiente para abordar una cuestión tan global. Como se aprecia, interoperabilidad y acceso abierto son, por consiguiente, dos conceptos muy ligados.

Así, por lo general, los impulsores del acceso abierto también potencian la interoperabilidad legal, de tal manera que las discrepancias entre las legislaciones de los distintos Estados miembros pueden convertirse en un obstáculo. De ello se desprende la necesidad de alinear las normas con la finalidad de que desaparezcan las barreras, limitaciones jurídicas. De hecho, algunos autores ven en la interoperabilidad legal una forma de combatir la fragmentación existente (Weber, 2014).

Organización Industrial y Gestión de Empresas de la Escuela de Ingeniería de la Universidad de Sevilla (Sebastián Lozano-Segura).

⁴⁷ Nota 11. Si dos o más sistemas se pueden comunicar entre sí, gozan de interoperabilidad sintáctica, lo que se logra con formatos de datos y protocolos y estándares específicos, tales como XML o SQL u otros de menor nivel.

Con anterioridad a la era del conocimiento, a la revolución digital, los regímenes legales quedaban enmarcados en un Estado, en un país. El punto de unión era la soberanía nacional. El antecedente de la interoperabilidad legal, probablemente, pueda encontrarse en el derecho internacional, en la medida en que este ha sido el instrumento usado hasta ahora para desarrollar la interoperabilidad legal.

Sin embargo, con la red parece que las fronteras han quedado difuminadas y el derecho internacional y las instituciones que en el mismo existen resultan insuficientes para satisfacer la interacción y la velocidad de la misma a día de hoy (Weber, 2014). De hecho, ya existen ejemplos en la actualidad de regímenes generales como el que afecta a la regulación de los nombres de dominio a nivel mundial.

Sin lugar a dudas, la interoperabilidad es una condición previa de la tecnología (Gasser y Palfrey, 2012). La interoperabilidad legal permite la circulación de mercancías, servicios e información dentro de sistemas legales (Gasser y Bukert, 2003).

Existen distintas formas de abordar la interoperabilidad legal, de arriba abajo o viceversa (Gasser y Palfrey, 2012). En los siguientes puntos, abordarán distintos ejemplos de cómo hacer frente a la interoperabilidad legal desde la fragmentación que aún existe. La precondition de la existencia o desarrollo de la interoperabilidad legal está basada en un principio básico que es el reconocimiento mutuo de los sistemas (Gasser y Palfrey, 2012).

El punto de partida tiene que ser, necesariamente, la definición de la interoperabilidad legal. Si la interoperabilidad en lo que respecta a la distinta información y tecnologías de comunicación permite que los datos se muevan entre distintos sistemas de información, en el caso de los sistemas legales puede decirse que son interoperables cuando el coste de la fragmentación regulatoria es lo bastante bajo para las personas, los bienes y los servicios, de acuerdo con la regulación, como para moverse con facilidad entre sí. Permite la diversidad jurisdiccional, evitando al mismo tiempo los efectos de bloquear a los sujetos en un sistema dado (Tréguer, 2013).

Otra definición de interoperabilidad legal de datos podemos encontrarla en una presentación de Uhler en la que se define de la siguiente forma: «los términos, condiciones y derechos legales de las bases de datos de dos o más fuentes son

compatibles y los datos debe combinarlos cualquier usuario sin comprometer los derechos legados de ninguna de las fuentes usadas». ⁴⁸

Prosigue Uhlir diciendo que la interoperabilidad legal es especialmente importante en la investigación pública, habida cuenta de que en ella se usan múltiples fuentes, se reutilizan, se combinan y vuelven a diseminarse.

En los últimos años, y más aún durante los últimos meses, una multiplicidad de grupos multinacionales y multifuncionales (abogados, ingenieros, científicos, libreros, etc.) están trabajando y publicando en relación con la interoperabilidad legal. De alguna forma, se está trabajando en la construcción de un nuevo marco jurídico ya no nacional ni comparado, sino global, que permita desdibujar los límites territoriales en lo que respecta a los intercambios científicos. Se hará después una reflexión sobre este tema aplicado a los datos de biodiversidad que, a fin de cuentas, es el objeto de esta tesis.

La noción de «interoperabilidad legal» (Alonso García, 2016b) se entiende hoy, aunque desde hace poco tiempo, de forma universal y desde una perspectiva amplia, como el lado jurídico o los aspectos jurídicos que necesitan tenerse en cuenta y solventarse siempre que se pretende lograr esquemas de «interoperabilidad técnica» de datos digitales. Esto es así a partir del reconocimiento de que:

El otro reto que tiene la interoperabilidad (global) de datos consiste en lograr la «interoperabilidad legal» que, de no lograrse, condiciona, si no impide, el procesado automático de datos basado en interacciones *machine-machine* llevados a cabo globalmente y, por tanto, la susceptibilidad de acceso y la reutilización de datos entre operadores situados en distintas jurisdicciones nacionales, que, como soberanas, regulan de manera diferente el derecho aplicable al carácter público o confidencial de los datos por distintos motivos, de interés privado o público, en cualquier caso de interés general normalmente, la responsabilidad de los errores o el mal uso de los datos o de las bases que los contienen, la protección del *software* que permite acceder o reutilizar los datos, los controles de su calidad o la propiedad misma de los datos y a los que controlan su uso a través de los de propiedad intelectual o industrial que les protegen» (Alonso García, 2016). ⁴⁹

Definir la interoperabilidad en abstracto no parece muy complicado. A la postre, la «interoperabilidad» consistirá en facilitar que los sistemas puedan interactuar y, por

⁴⁸ Definición extraída de *Data Sharing Task Force* (pendiente de publicación). Borrador del *White Paper on Legal Options for the Exchange of Data through the GEOSS Data*. Presentación disponible en http://www.nsgic.org/public_resources/02_Uhlir_Legal-Interoperability-of-Data_NSGIC-Conf_Feb13.pdf. (Consultado por última vez el 23 de abril de 2016).

⁴⁹ Extraído de *Principles on the Legal Interoperability of Research Data Implementation Guidelines*. Borrador de 9 de febrero de 2016.

lo tanto, «interoperabilidad legal» significará facilitar que los sistemas jurídicos no limiten o limiten lo mínimo posible la circulación de la investigación científica.

La interoperabilidad legal resuelve cómo llegar al origen, pero eso no es suficiente. Consiste, por lo tanto, en informar al que va a usar el dato de las condiciones jurídicas con las que se puede usar el dato: puede usarse, descargarse, etc. Pero además, a la hora de mezclar datos lo que se plantea es cómo se mezclan datos sujetos a distintas condiciones jurídicas.

Esto ya plantea problemas diferentes, que habrá que tener en cuenta a la hora de hablar de interoperabilidad legal o de cuáles son los límites que esta debería tener. En este momento, en el marco de la RDA⁵⁰ un grupo —del que forma parte integrante la autora— está analizando y trabajando sobre la construcción de unos principios para la interoperabilidad legal y sobre lo que es necesario tener en cuenta en este marco. Se abordarán un poco más adelante. En el análisis previo realizado para la elaboración de dichos principios, se analiza cuándo se da la interoperabilidad legal.

En efecto, la interoperabilidad legal sucede cuando concurren los siguientes presupuestos:

- Que las condiciones de uso sean claras y puedan ser determinadas en relación con cada uno de los conjuntos de datos.
- Que las condiciones de uso que se imponen a cada uno de los conjuntos de datos permitan la creación y reutilización de productos combinados o derivados.
- Que los usuarios puedan tener derecho de acceso y de uso a cada uno de los conjuntos de datos sin necesidad de pedir la autorización de los titulares de los datos caso por caso, asumiendo que las condiciones acumuladas para el uso de cada uno de todos los conjuntos de datos se cumplen.

Además, la interoperabilidad legal también implica que la búsqueda de permisos y licencias para el uso de los datos también se producirá en el entorno de la red. Cuando se combinen datos de una multiplicidad de fuentes, los conjuntos de datos resultantes deberán incorporar la acumulación de restricciones impuestas por todas y cada una de las fuentes.

⁵⁰ *Principles on the Legal Interoperability of Research Data. Implementation Guidelines*. Borrador de 24 de julio de 2016. *Definitions and Discussion of Terms. Interoperability of Data*, aún pendiente de publicación.

Por consiguiente, tiene que hacerse un seguimiento de cada una de las restricciones aplicables. Cuantas menos restricciones se apliquen a los conjuntos de datos, menos restricciones tendrán los conjuntos de datos derivados de estos. La forma más sencilla para que exista la interoperabilidad legal es que los conjuntos de datos no tengan ningún tipo de restricción jurídica. Es fundamental que los científicos conozcan los retos y oportunidades resultantes de las normas de propiedad intelectual e industrial.

La finalidad de estos principios es contribuir a lograr la interoperabilidad legal de los datos resultantes de las investigaciones desde un punto de vista nacional, de disciplina y de sector. Una de las recomendaciones fundamentales es que tanto los investigadores como los grupos de investigadores e instituciones pongan sus conclusiones e investigaciones en el dominio público y, por defecto, recurriendo al uso de las licencias adecuadas.

Más allá de todo esto, lo cierto es que los datos producidos con recursos públicos deben ser públicos y de libre acceso y reutilización para todo el mundo; este es el nuevo paradigma (Alonso García, 2016b).

Como se ha visto en puntos anteriores, los científicos no han interiorizado ese nuevo paradigma, aunque están proliferando los grupos de estudio en los que se trabaja para concienciar a estos científicos con el objetivo de que reduzcan las barreras de acceso a estos datos. La interoperabilidad legal, más allá de normas que tardarán más en darse, requiere un cambio cultural, un cambio de paradigma en los «titulares» de los datos que los pongan a disposición de los demás de forma abierta (en el dominio público o mediante el uso de licencias abiertas).

Es importante, a los efectos de la interoperabilidad legal, tener muy en cuenta las restricciones. La clave es poner el mínimo de restricciones posibles para la reutilización de los datos, puesto que, si se ponen barreras en cada reutilización de los mismos, se irá restringiendo el uso de forma paulatina.

Otra herramienta para lograr el mencionado objetivo son las políticas de OA establecidas a nivel de los gobiernos (National Science Foundation, Comisión Europea...), de manera que el acceso a las subvenciones tenga como requisito poner los datos a disposición de todos sin establecer restricciones de uso. El cambio de paradigma afecta también a la estructura de la sociedad. En efecto, se ha producido un cambio de la cultura competitiva a la cultura colaborativa (Scharmer, 2016). La colaboración hoy es una manera de crecer y llegar mucho más lejos. Hoy en día, que uno crezca no significa necesariamente que otro no lo haga (Beck y Cowan, 1996). Se está haciendo una transición cultural de un patrón de comportamiento de apropiación de las cosas para uso exclusivo a una búsqueda de propósitos comunes en la que haya muchos focos de entrada y salida de datos,

donde estos se reutilicen (idealmente) sin restricciones. Por lo tanto, parece que la interoperabilidad legal, más allá del cambio normativo, requiere un cambio de enfoque en lo que a la gestión de los datos se refiere. En el siguiente capítulo se analizarán ejemplos, casos concretos.

Es clave, por lo tanto, ir desarrollando ámbitos globales que permitan el intercambio de datos (Arzberger *et al.*, 2004). Ya algunos autores como Oxenham (2016) reflexionan sobre el hecho de que la poca claridad jurídica y la confusión que existe sean una amenaza para el intercambio de datos científicos. La creciente preocupación por incumplir las normas de las distintas jurisdicciones, así como la imposibilidad de saber a ciencia cierta las condiciones de uso y reutilización de los datos pueden resultar disuasorias a la hora de hacer uso de los mismos.

La necesidad de pedir permiso para el uso de cualquier dato por anticipado, incluso aunque se pudiera identificar al titular del dato, complicaría sobremanera el uso o la reutilización de los mismos, lo que sería casi imposible o poco práctico.

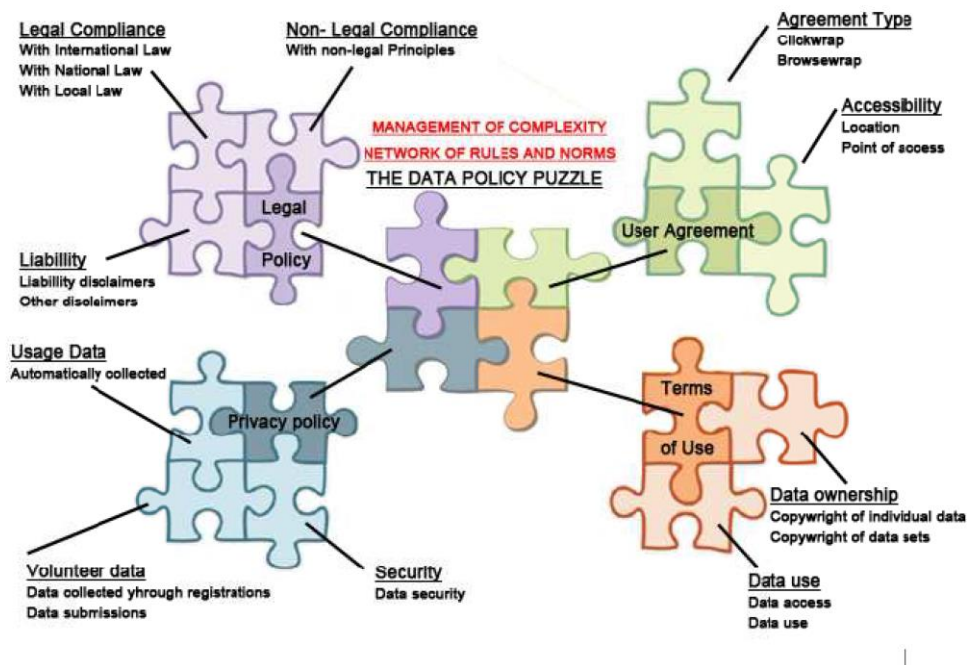


Figura 47. Políticas de datos para la participación pública en la investigación científica elaborado por Anne Bowser. Fuente: Alonso García, 2016b.

Alonso García en este mismo artículo se plantea si están armonizadas las exigencias de la interoperabilidad legal.

4.3.1. *Las posibles alternativas al copyright: el nacimiento del movimiento copyleft*

En los últimos años han ido proliferando instrumentos que facilitan a los titulares o propietarios de los datos «abrir» su uso y su reutilización a otras personas. Es una forma de contribuir a la simplificación y mejor comprensión de la necesidad de interoperabilidad legal. Para ello se han creado distintos sistemas de licencias que permiten simplificar el uso del *copyright* y los derechos de uso y reutilización de los datos.

Esta construcción de alternativas está ligada al cambio de paradigma en la era del conocimiento en la que se están enfrentando, como se ha ido analizando a lo largo de este trabajo, dos fuerzas opuestas: la libertad y la apertura frente a la clausura y el control (Vercelli, 2004).

Existe una tendencia de todo punto incorrecta a pensar que puede disponerse libremente de todo lo que aparece en Internet (Broussard, 2007). De hecho, algunos autores como Lessig (2001 y 2004a) ponen de manifiesto que cada vez que se copia algo de Internet se están vulnerando derechos de autor, puesto que no se tiene permiso para hacer uso de ello.

La finalidad del movimiento de *copyleft* (Dusollier, 2004) no es otro que poner el máximo de información en el dominio público, facilitando su reutilización. Ese término se acuña por contraposición al *copyright*, aunque no es un antónimo como tal, ya que mientras *copyright* significa «todos los derechos reservados», *copyleft* quiere decir «algunos derechos reservados» (Brown, 2010). De hecho, el concepto de *copyleft* nace en el ámbito del *software* como una forma de hacer programas abiertos que puedan ser enriquecidos por cualquiera. No es este el objeto del presente trabajo, pero sí lo es, en cambio, uno de los ejemplos principales de *copyleft*: las licencias Creative Commons.

4.3.2. *Ejemplos concretos del movimiento copyleft: de las GNU a las licencias Creative Commons*

El movimiento *copyleft* apareció a finales de 2002. Su fuente de inspiración es la fundación de *free software* GNU de licencia al público general (*GNU General Public License*). Por lo tanto, nace en Estados Unidos. A pesar de que la propiedad intelectual es una institución moderna y la creación humana es anterior al *copyright*, ambas están muy ligadas, lo que se acentúa con la tecnología (Xalabarder Plantada, 2006). De alguna manera, el cambio tecnológico ha hecho que el derecho de propiedad intelectual o *copyright* pase de ser un monopolio de explotación a ser un derecho a controlar el uso (Xalabarder Plantada, 2006), como ha quedado varias veces puesto de manifiesto en este trabajo.

Las licencias de Creative Commons (Creative Commons, s. f.) son una nueva manifestación de la tensión entre apertura y control. Se configuran como un movimiento social con la finalidad de crear una plataforma para un amplio conjunto de personas que tienen interés en desarrollar el acceso a los distintos trabajos (Elkin-Koren, 2005). En Estados Unidos ya no existen requisitos formales para que exista *copyright*, sino que el creador de una obra no necesita ningún trámite formal para protegerla, lo que ha reducido el número de obras que se encuentran en el dominio público.

Como respuesta a ello surgen este conjunto de instrumentos jurídicos con modelos de contratos de licencias que pueden utilizarse a través de una plataforma web que hace posible licenciar los trabajos al público con algunos tipos de restricciones. Creative Commons introduce (Elkin-Koren, 2005) una forma innovadora de usar los derechos con la finalidad de facilitar un cambio social. De alguna forma, abre una vía intermedia entre los defensores a ultranza del *copyright* y sus detractores, partidarios del *public domain*. Reconoce la existencia de derechos de propiedad intelectual e incentiva a los autores para que sean los que compartan con el mínimo de restricciones posibles. La idea es hacer más accesibles y reutilizables los datos, minimizando los costes de acceso y potenciando, de esta manera, el acceso a los trabajos creativos.

La ideología que subyace puede sintetizarse de la siguiente forma (Elkin-Koren, 2005):

- La creatividad depende de la posibilidad de acceso y utilización de trabajos preexistentes.
- Las regulaciones del *copyright* crean nuevas barreras de acceso a los trabajos, por lo que se convierten en un obstáculo para compartir y reutilizar los trabajos creativos.
- Los elevados costes asociados al régimen de *copyright* limitan la posibilidad de que los individuos accedan y reutilicen trabajos creativos.
- Los derechos de propiedad intelectual tienen que usarse de forma que promuevan el intercambio (compartir) y la reutilización.

Este sistema estandariza y facilita su uso y puede utilizarse en más de cincuenta jurisdicciones diferentes. Todas las licencias se expresan en tres niveles de lectura:

- Un primer nivel de lectura humana, iconográfica, para facilitar la identificación, que describe los usos que autoriza el autor. Es el nivel más visible y lo que caracteriza a estas licencias.

- Un segundo nivel de lectura jurídica, que no deja de ser humana. Sería la licencia que redactaría un abogado o un profesional de propiedad intelectual o *copyright*.
- Un tercer nivel de lectura tecnológica, que solamente pueden interpretar las máquinas (los programas) y, en concreto, los motores que permiten identificar las obras que tengan una licencia de estas características (Xalabarder Plantada, 2006).

Como se explica en el propio sitio web de Creative Commons, el hecho de aplicar este tipo de licencias no significa que la obra no tenga *copyright*. Son licencias que ofrecen algún tipo de derechos con sujeción a algunas condiciones que se presentan a continuación.



Reconocimiento (*Attribution*): En cualquier explotación de la obra autorizada por la licencia hará falta reconocer la autoría.



No comercial (*Non commercial*): La explotación de la obra queda limitada a usos no comerciales.



Sin obras derivadas (*No Derivate Works*): La autorización para explotar la obra no incluye la transformación para crear una obra derivada.



Compartir igual (*Share alike*): La explotación autorizada incluye la creación de obras derivadas siempre que mantengan la misma licencia al ser divulgadas.

Con estas condiciones se pueden generar las seis combinaciones que producen las licencias Creative Commons:



Reconocimiento (*by*): Se permite cualquier explotación de la obra, incluyendo una finalidad comercial, así como la creación de obras derivadas, la distribución de las cuales también está permitida sin ninguna restricción.



Reconocimiento-NoComercial (*by-nc*): Se permite la generación de obras derivadas siempre que no se haga un uso comercial. Tampoco se puede utilizar la obra original con finalidades comerciales.



Reconocimiento-NoComercial-CompartirIgual (*by-nc-sa*): No se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.



Reconocimiento-NoComercial-SinObrasDerivadas (*by-nc-nd*): No se permite un uso comercial de la obra original ni la generación de obras derivadas.



Reconocimiento-CompartirIgual (*by-sa*): Se permite el uso comercial de la obra y de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.



Reconocimiento-SinObraDerivada (*by-nd*): Se permite el uso comercial de la obra pero no la generación de obras derivadas.

Además de todas estas combinaciones, en el año 2009 se crea la licencia CC0, pensada fundamentalmente en la ciencia. La finalidad de esta licencia es renunciar a todos los derechos de autor y derechos conexos en relación con un trabajo. No puede usarse sin el permiso del autor. La propia organización de Creative Commons recomienda no usar esta licencia para trabajos que ya estén en el dominio público.

El uso de la licencia CC0 significa que existen derechos de autor y que el propietario renuncia a ellos. En algunos casos —véase el ejemplo de la NASA explicado abajo— no es posible usar esta licencia porque no existen derechos de autor, sino que los datos por definición están en el dominio público.

Como ejemplo de esto, la NASA lideró un caso de estudio en el Antártico (Anon., 1959).⁵¹ La elaboración del caso y los resultados del mismo se hicieron usando una licencia denominada CC0 (Creative Commons, s. f.b), que se estudiará con más detalle en el marco de contextualización de la interoperabilidad; se configura como un *waiver*, de alguna manera, una renuncia al *copyright* sobre los resultados.

Sin embargo, la NASA se negó a incorporar los resultados a sus estudios puesto que, por definición, todos los datos de la NASA están en el dominio público y no se puede imponer ninguna restricción sobre los mismos. Los datos generados en el marco del Tratado Antártico no son los únicos no susceptibles de protegerse con *copyright*. Aunque el efecto de la licencia sea precisamente «ceder» todos los derechos, la discusión está en si existen derechos susceptibles de protección o no; si los derechos no existen, no pueden licenciarse, así que los de la NASA tampoco.

Por lo tanto, la condición de uso de este tipo de licencias es precisamente la preexistencia, el reconocimiento de que existen derechos de autor (Xalabarder Plantada, 2006). El uso de este tipo de licencias exige, en todo caso, reconocer al autor de la obra.

Si bien se ha recomendado el uso de estas licencias con el mínimo de restricciones posibles, no debe olvidarse que el uso de licencias de cualquier índole restringe el

⁵¹ Todas las investigaciones de la Antártida están sujetos al Tratado del Antártico. Los datos son de dominio público.

uso de los datos y que la combinación de varias licencias puede limitar los usos y reutilizaciones de los datos de forma innecesaria. En efecto, en el caso de que se usen varias protecciones, siempre prevalecerá la más restrictiva: lo que los autores han llamado el *lowest common denominator* (mínimo denominador común).

Pone de manifiesto que la aplicación de licencias a los datos científicos dificulta el acceso a los datos (Stodden, 2009). Desaconseja, sin duda, la aplicación de *share alike* por la confusión que introduce en la reutilización de los datos. En efecto, este tipo de licencia pone muchas trabas a los datos.

Las comunidades científicas, en muchos casos, recomiendan el uso de este tipo de herramientas con la finalidad de potenciar el acceso a los trabajos, datos, etc.

4.3.3. Distintas propuestas para ir construyendo una interoperabilidad legal ¿global?

Los partidarios del movimiento de ciencia abierta defienden que, para que la ciencia pueda funcionar de forma adecuada y existan beneficios para la totalidad de la humanidad, es fundamental garantizar el acceso abierto a los detalles de la práctica científica, así como el acceso abierto a las revistas científicas, a los datos abiertos, a las fuentes y códigos científicos y garantizar un acceso a las publicaciones científicas y a los datos en los que se basan (Peretti Pezzi, 2015).

Como consecuencia de todo este movimiento, se empiezan a construir una serie de principios en distintos lugares, con la finalidad de que los datos científicos estén cada vez más disponibles y puedan reutilizarse. De este modo se eliminan restricciones y barreras de uso; a fin de cuentas, se construye interoperabilidad legal.

4.3.3.1. Los principios y directrices de implementación de la Research Data Alliance

La Research Data Alliance (en adelante, RDA) (Research Data Alliance, s. f.) construye los puentes sociales y técnicos que hacen posible el intercambio de datos en abierto. Nació como organización comunitaria en 2013, creada por la Comisión Europea, la National Science Foundation, el National Institute of Standards and Technology de Estados Unidos y el Government's Department of Innovation de Australia con el objetivo de construir la infraestructura técnica para permitir el intercambio abierto de datos.

La visión de la RDA es facilitar un ecosistema para facilitar la creación de un entorno de apertura en el que los investigadores e innovadores compartan abiertamente los datos entre tecnologías, disciplinas y países para abordar el gran reto al que se enfrenta la sociedad.

El actual marco global de datos de investigación está muy fragmentado, por disciplinas o por ámbitos: oceanográfico, relacionado con las ciencias de la vida y de la salud, de la agricultura, espacial, del clima, etc. Cuando se trata de actividades interdisciplinarias, las nociones de «bloques de construcción» de las infraestructuras de datos comunes y la construcción de «puentes de datos» específicos se están convirtiendo en metáforas aceptadas para abordar la complejidad de los datos y permitir el intercambio de datos.

Como ponen de manifiesto Natalia Manola y Andrés Wiebe en una motivación que preparan para la RDA, es necesario construir un grupo de trabajo dentro de este marco con la finalidad de apoyar la interoperabilidad. A fin de cuentas, es necesario crear marcos de propiedad intelectual compartida donde la información sea interoperable y se generen de esta forma marcos ricos en conocimiento que puedan usarse de forma segura.

Prosiguen, en la línea de otros autores, considerando que los datos de investigación no pueden protegerse mediante los sistemas de propiedad intelectual o *copyright*, pues, de primeras, estos datos no son susceptibles de quedar protegidos. Podrán quedar protegidos en los supuestos en los que estén ordenados en bases de datos (tanto en Estados Unidos como en Europa, con los condicionantes analizados en puntos anteriores). Es raro que las bases de datos con datos de investigación puedan resultar protegidas. Los autores propusieron la creación de un grupo de trabajo a estos efectos.

La RDA permite compartir los datos sin barreras a través de grupos de trabajo y grupos de interés formados por expertos de todo el mundo —académicos, industrias y Gobiernos—. La participación en RDA está abierta a cualquier persona que esté de acuerdo con sus principios rectores de apertura, consenso, equilibrio y armonización, así como con su enfoque comunitario y sin ánimo de lucro.

Con más de seis mil setecientos miembros de ciento treinta y seis países (en marzo de 2018), la RDA proporciona un espacio neutral donde sus miembros pueden reunirse a través de grupos de trabajo e interés globales para desarrollar y adoptar la infraestructura que promueva el intercambio de datos y la investigación impulsada por datos, así como acelerar el crecimiento de una comunidad de datos cohesionados que integre contribuyentes a través del dominio, investigación, fronteras nacionales, geográficas y generacionales.

En esta línea, se crea un grupo de interés en temas de interoperabilidad legal (Research Data Alliance, s. f., RDA/CODATA Legal Interoperability IG) donde se encuentran representados científicos y abogados de varios lugares del mundo, integrados en distintos tipos de organizaciones. La finalidad de este grupo será abordar y promover los siguientes objetivos:

- Definir la interoperabilidad de los datos de investigación y articular lo importante.
- Documentar y analizar hasta cuatro casos de estudio en las áreas de geociencia, investigación de biodiversidad, ciencias sociales y humanidades en relación con soluciones de interoperabilidad legal en contextos internacionales y multidisciplinarios.
- Desarrollar y publicar los principios *core* y directrices de mejores prácticas a través de las cuales se puede alcanzar la interoperabilidad legal.
- Trabajar con los grupos clave de *stakeholders* para alcanzar los principios *core* y las directrices de mejores prácticas adoptadas.
- Con carácter general, promover una mejor comprensión y un mayor uso por parte de los grupos clave de *stakeholders* dentro de la comunidad científica de los enfoques acordados en materia de interoperabilidad de los datos de investigación, con la finalidad de subrayar y potenciar una mejor integración y reutilización de esos datos.

Para ello, deciden elaborar una propuesta de principios sencillos que contribuyan a mejorar la interoperabilidad legal de los datos. Los distintos borradores elaborados se han ido presentando a organizaciones y otros científicos con la finalidad de testar si dichos principios daban respuesta o, de alguna forma, eran útiles para la consecución de su objetivo: contribuir a una interoperabilidad legal más global.

Publicaron los Principios de Interoperabilidad de Datos de Investigación (RDA-CODATA Legal Interoperability Interest Group, 2016) acompañados de unas directrices de interpretación que pretenden aportar luz en su aplicación. La construcción de los principios se ha hecho de forma sinérgica, es decir, su aplicación es mejor cuando interactúan entre sí. Lo que subyace en la elaboración de estos principios es el trabajo o fomentar entornos en los que los datos estén en abierto, sin restricciones o con las mínimas posibles. Por lo tanto, el enfoque no es propiamente la interoperabilidad legal. Se definen a continuación los principios mencionados.

Primero. Facilitar el acceso legal y la reutilización de los datos de investigación

La finalidad de este principio es convertir los datos abiertos en la regla y, en su defecto, que se apliquen el mínimo de restricciones posibles. Se recogen una serie de recomendaciones de uso para lograr esta finalidad.

En efecto, los Gobiernos, instituciones e investigadores tienen la posibilidad de poner los datos de investigación en el dominio público (sobre el tema del dominio público se volverá más adelante, para hacer un análisis más profundo del mismo) sin restricciones de uso. Se elabora una tabla con los posibles mecanismos para lograr

datos abiertos y conseguir la interoperabilidad legal que queda reproducida a continuación (Doldirina *et al.*, en prensa):

RDA-CODATA Legal Interoperability Interest Group 2016

Table 1: Summary of Legal Mechanisms that Promote Open Data and Legal Interoperability¹

Type of Legal Mechanism	Summary Description
Governmental Mechanisms	
Intergovernmental Agreements	Governments can enter into treaties or international agreements (multilateral or bilateral) that create binding obligations among governments to exclude government generated or funded data from intellectual property rights protection, and should be made publicly available if no other restrictions apply.
Intergovernmental Policies	International or intergovernmental organizations can adopt policies to make certain types of publicly funded research data publicly available without restriction on use or reuse.
National Legislation, Regulations, or Policies	National governments can enact laws, issue regulations at the ministerial level, or adopt a broad range of policies to place publicly funded research data in the public domain. Government research funding agencies and government data centers or services can include requirements in grants or contracts to make resulting research data publicly available without restrictions on use or reuse.
Nongovernmental and Governmental Mechanisms	
Creative Commons Public Domain Mark (CC PDM, http://creativecommons.org/chouse/mark/)	The CC Public Domain Mark is used to mark collections of data over which any potential copyright has expired or never existed, and thus are already in the public domain, enabling their more ready identification in global web searches. Except for data arising in the public domain (e.g., facts) and data in historic documents that have been digitized, few collections of data should likely have this mark applied.
Creative Commons No Rights Reserved Instrument (CC0, http://creativecommons.org/chouse/zero/)	To the extent possible under law across the world, the person or authority that associates CC0 with the work waives all copyright and related or neighboring rights to the work, such as the E.U. <i>sui generis</i> database right. It has been declared compatible with the Open Definition by the Open Definition Advisory Council http://opendefinition.org/licenses/
Open Data Commons Public Domain Dedication and License (PDDL, http://www.opendatacommons.org/licenses/pddl/1-0/)	The PDDL allows the database user to “copy, distribute and use the database”; “produce works from the database”; and “modify, transfer and build upon the database.” It has been declared compatible with the Open Definition by the Open Definition Advisory Council http://opendefinition.org/licenses/
Creative Commons Attribution 4.0 License (CC-BY 4.0, https://creativecommons.org/licenses/by/4.0/legalcode)	The CC-BY 4.0, is the least restrictive of the Creative Commons licenses. It permits the user to share and adapt material with minimal attribution requirements. It has been declared compatible with the Open Definition by the Open Definition Advisory Council http://opendefinition.org/licenses/
Nongovernmental Policies and Agreements	Nongovernmental organizations can implement policies to retain no rights in research data or conclude agreements among institutional members of a consortium to retain no rights in research data.

Figura 48. Resumen de los mecanismos legales que promueven el acceso abierto y la interoperabilidad legal. Fuente: Doldirina, Eisentadt, Onsrud y Uhler (pendiente de publicación).

Sobre el análisis de los mecanismos mencionados, se volverá más adelante en profundidad.

Segundo. Determinar los derechos y responsabilidades de los datos

Aquellas personas que contribuyan a la difusión de los datos de investigación tienen que establecer quién es el titular de los datos antes de proceder a su difusión.

Además, deben asegurarse de que se respetan las normas en las distintas jurisdicciones.

Otra de las recomendaciones de esta directriz es animar a los expertos en estos temas a colaborar a la construcción de las normas y leyes en las distintas jurisdicciones. Por último, conseguir que exista una buena formación en los investigadores en relación con los derechos y responsabilidad derivados de los datos.

Tercero. Equilibrar los intereses jurídicos

En lo que respecta a este punto, la clave está en promover *waivers* o renunciaciones a los derechos de propiedad intelectual o *copyright* y que se distribuyan con el mínimo de restricciones posibles en las colecciones financiadas con fondos gubernamentales. Es más, se pide que tanto los Gobiernos como los que investiguen con financiación pública justifiquen las restricciones que incluyan en los datos de investigación.

Se recomienda, además, el desarrollo de políticas y normas que promuevan el acceso público a los datos de investigación, así como la necesidad de reducir los periodos de embargo lo máximo posible. Se recomienda, también, limitar cuanto se pueda la introducción de restricciones mediante licencias o acuerdos en el caso de datos de investigación procedentes de financiación pública.

Cuarto. Declarar los derechos de manera transparente y clara

En la medida de lo posible, se recomienda el uso de herramientas electrónicas y estandarizadas a la hora de abordar la retención de derechos, si es que esta procede.

Se recomienda, asimismo, buscar asistencia jurídica para entender la aplicación en las distintas jurisdicciones y que los titulares de los derechos den la máxima información posible en relación con las particularidades que sean de aplicación y las condiciones de uso.

Quinto. Promover la armonización de los derechos en los datos de investigación

Existen distintas formas de abordar la armonización de derechos en lo que respecta a los datos de investigación. Por una parte, puede recurrirse a herramientas de carácter voluntario como pueden ser las licencias (ya analizadas en el contexto del *copyleft*). La armonización también puede realizarse a través de instrumentos jurídicos como los tratados internacionales.

Sexto. Proporcionar la atribución y el crédito apropiados para los datos de la investigación

La norma de atribución debería ser una convención resultante de las buenas prácticas en el ámbito de los datos de investigación y no un requerimiento legal.

La finalidad de la elaboración de estos principios es facilitar la reutilización de datos y estos son de seguimiento voluntario y están sujetos a continuas revisiones y actualizaciones con la finalidad de ir dando respuesta a las cuestiones que surjan.

De alguna manera, también su contribución se encuentra en acercar las visiones científicas de los «propietarios» o no de los datos de investigación a la visión jurídica. En realidad, es una forma de que los científicos entiendan también que una correcta gestión de los datos, un intercambio y archivo de los mismos pueden contribuir a una ciencia más abierta.

A fin de cuentas, son unos principios nacidos en el marco de una organización orientada a la *open science*, al intercambio de datos y, por lo tanto, tratan de que los propios generadores de los datos establezcan cada vez menos restricciones. Es una manifestación más de la necesidad de orientarse hacia la elaboración de un marco global de interoperabilidad.

Si bien es cierto que se han multiplicado las iniciativas para ir promoviendo la accesibilidad y reutilización de los datos, existen todavía obstáculos legales que deben superarse (Contreras y Reichman, 2015).

El trabajo de este grupo de interoperabilidad legal no ha terminado en el marco de la RDA. Varios artículos sirvieron también para apoyar estos principios: uno, relacionado con las consecuencias del intercambio de datos que es fuente generadora de responsabilidades compartidas (Minster y Clement, 2016); otro, relacionado con las razones sobre las que se fundamenta la necesidad de que exista la interoperabilidad legal de los datos científicos (Chen, 2016) y otro, relacionado con las cuestiones que se plantean en relación con la interoperabilidad de los datos científicos (Uhlir *et al.*, 2016). El valor de los datos abiertos es incuestionable (Uhlir, 2015).

Con el desarrollo del *big data* y de la e-ciencia, se ha acentuado el énfasis en la accesibilidad y reutilización de los datos con la finalidad de mejorar la investigación y sus aplicaciones (Uhlir, 2016). Por ello, facilitar la interoperabilidad legal (la más relevante a los efectos de este trabajo), pero también la técnica y semántica, es clave para la integración de los datos y la maximización de su valor. Uhlir (2016) continúa analizando *use cases* con la finalidad de evaluar el impacto de los principios y de su aplicación en la práctica para poder irlos corrigiendo y modulando.

4.3.3.2. Los principios del Foro de Belmont

El Foro de Belmont tiene como objetivo acelerar la entrega de la investigación medioambiental necesaria para eliminar las barreras críticas a la sostenibilidad al alinear y movilizar los recursos internacionales, agregando valor a las inversiones

nacionales existentes y alentando las asociaciones interdisciplinarias y transdisciplinarias. Creado en 2009, representa a un grupo de los principales financiadores mundiales y emergentes de la investigación del cambio medioambiental global. Está integrado por miembros de Austria, Australia, Brasil, China, la Comisión Europea, Francia, Alemania, India, Italia, Japón, Noruega, Qatar, Sudáfrica, Suecia, Taiwán, Reino Unido y Estados Unidos. Además, los organismos de financiación de Nueva Zelanda, la Federación de Rusia y Suecia expresaron su interés de unirse al Foro de Belmont como nuevos miembros. El Consejo Internacional de Ciencias Sociales (ISSC), el Consejo Internacional de Ciencias (ICSU), el Instituto Internacional de Análisis de Sistemas Aplicados (IIASA) y el Programa de Investigación sobre Cambio Global de Estados Unidos (USGCRP) son socios (no financiadores) (Comisión Europea, s. f., Belmont Forum).

Es un grupo de representantes de alto nivel de las principales agencias a escala mundial (Schmidt *et al.*, 2016) que coordina la subvención de acciones de investigación colaborativa con la finalidad de abordar los retos y oportunidades del cambio medioambiental global. Ha creado, por lo tanto, una plataforma global de cooperación y coordinación que comparte una visión (Belmont Forum, 2016).

Los copresidentes actuales son, en primer lugar y como representante de la Comisión Europea, Kurt Vandenberghe, copresidente desde 2015, hasta hace poco director de la acción sobre el clima y la eficiencia de los recursos y ahora director de desarrollo y coordinación de políticas, en la DG de Investigación e Innovación de la Comisión Europea. El otro copresidente y representante de la Fundación de Investigación de Sao Paulo (FAPESP) es Gilberto Câmara, director general que sucederá a la agencia de Japón como segundo copresidente.

Las iniciativas transnacionales de investigación promovidas por el Foro de Belmont se desarrollan en sinergia con Horizonte 2020 y con proyectos de innovación desarrollados por los Estados miembros de la Unión Europea. La participación de la Dirección General de Investigación e Innovación en el Foro de Belmont es testigo de la convicción de la Comisión Europea de que abordar las crecientes necesidades derivadas de los desafíos mundiales del cambio debe ser respaldado por una cooperación científica internacional abierta, transdisciplinaria y orientada a los resultados.

Refuerza la ya importante participación financiera de la Unión Europea en iniciativas internacionales como el Sistema Mundial de Sistemas de Observación de la Tierra, la Alianza Global para las Enfermedades Crónicas, la Asociación de Investigación e Innovación de la Unión Europea-África sobre Seguridad Alimentaria y Nutricional y Agricultura Sostenible y el Futuro Tierra. De este modo, se consolida aún más el lugar en el que se sitúa la investigación financiada por la

Unión Europea en el centro de la actual tendencia mundial hacia una participación más activa de la ciencia, la tecnología y la innovación en la realización de los Objetivos de Desarrollo Sostenible después de 2015.

Como una de sus principales actividades, el Foro Belmont lanza Acciones de Investigación Colaborativa (AIC) sobre temas específicos. Las AIC tienen objetivos centralizados, pero la financiación de los componentes y proyectos individuales es responsabilidad de las distintas agencias participantes. Los programas financiados por la UE, como ERA-NET y las Iniciativas de Programación Conjunta de la Unión Europea, participan directamente en la aplicación de las AIC en los ámbitos siguientes: vulnerabilidad costera (2012), seguridad del agua dulce (2012), seguridad alimentaria (2013), e-infraestructura (2013), biodiversidad y ecosistema (2014), ártico (2014), montaña (2015), predictibilidad climática (2015). Las nuevas AIC en los ámbitos de urbanización global, transformación a la sostenibilidad y ‘acceso abierto a datos y e-infraestructuras’ han arrancado en 2016. Las dos primeras han sido denominadas como «Agencias AIC de Future Forum del Belmont Forum» porque contribuyen directamente a la consecución de los objetivos de la Tierra del futuro.

Por lo tanto, el foro de Belmont es totalmente complementario de Future Earth, una ambiciosa iniciativa nueva, lanzada en la cumbre Rio+20 tras la promoción de la Alianza de Ciencia y Tecnología para la Sostenibilidad Global (en suma, la alianza, compuesta por el Consejo Internacional de Ciencia, el Consejo Internacional de Ciencias Sociales, el Foro de Belmont, el PNUMA, la UNESCO, la Universidad de las Naciones Unidas y la Organización Meteorológica Mundial). Future Earth forma una plataforma global e independiente para la colaboración científica en la investigación sobre el cambio global y la sostenibilidad en torno a ocho desafíos claves de sostenibilidad identificados en Future Earth Vision 2025, a saber: «*Healthy and resilient cities*», «salvaguardar los bienes naturales», «descarbonización de los sistemas energéticos», «agua, energía, nexo de alimentación», «resiliencia, umbral y transformación», «producción y consumo sostenibles», «salud humana y planetaria» y «futuros rurales». Actualmente, se están explorando formas de colaboraciones e interacciones concretas entre Future Earth y el Foro de Belmont (incluida la Comisión Europea), lo que puede dar lugar a la creación de otras agencias en áreas de interés específico para Future Earth.

Con la necesidad de hacer frente a las cuestiones que se plantean en el siglo XXI nace el Belmont Challenge (Foro Belmont [Annual Meeting], 2016) cuya función puede resumirse en servir de apoyo para la investigación transdisciplinaria internacional de manera que proporcione el conocimiento necesario para comprender, mitigar y adaptarse al cambio medioambiental mundial.

4.3.3.3. Los *Principles for open data in science* o Principios Panton

Este concepto fue acuñado por Murray-Rust *et al.* (19 febrero de 2010) y recoge que los datos son la base de la ciencia: recabarlos, analizarlos, publicarlos, reanalizarlos, cuestionarlos y reutilizarlos. Sin embargo, existe toda una serie de barreras que no permiten acceder a los datos, que suponen restricciones de uso que aplican los editores o los suministradores de datos y que hacen que los datos sean difíciles de reutilizar. A esto hay que sumarle, como ya se ha puesto de manifiesto en una multiplicidad de ocasiones, la reticencia cultural a publicar los datos abiertamente (Molloy, 2011). Los investigadores temen poner «sus» datos a disposición de todos y no controlar lo que se hace con ellos o que no se les reconozca que son sus descubridores.

Como respuesta a todas estas cuestiones, surgen una serie de movimientos y de organizaciones que persiguen el acceso abierto como una forma de conseguir un mayor progreso en la ciencia. Una de estas organizaciones es Working Group on Open data in Science (el Grupo de Trabajo de Datos Abiertos en la Ciencia) de la Open Knowledge Foundation (Fundación de Conocimiento Abierto; en adelante, OKF por sus siglas en inglés).

La OKF es una organización que promueve el conocimiento abierto; dentro de esta idea, incluye los conceptos de datos abiertos, cultura libre, dominio público y otras áreas de conocimiento comunes (Open Knowledge Foundation, s. f.). Se creó en 2004 y se ha desarrollado como una red internacional de comunidades que desarrollan herramientas, aplicaciones y directrices que permiten abrir los datos y, como consecuencia de ello, descubrir y usar los datos. Trabajan en distintos ámbitos, todos unidos bajo el mismo paraguas de valores y principios, donde se comparte un mismo enfoque de apertura, conforme a la definición que da la OKF en la red y que, en resumen, dice que el conocimiento está abierto si cualquiera puede acceder a él, usarlo, modificarlo y compartirlo libremente sujeto, como mucho, a medidas que preservan la procedencia y la apertura (Open Knowledge International, s. f.).

En esencia, esta definición coincide con el de «abierto» con respecto al *software* como en la definición de *open source* y es sinónimo de *free* o «libre» como en la definición de *software* libre y en la definición de obras culturales libres.

- El término *trabajo* se utilizará para designar el elemento o pieza de conocimiento que se transfiere.
- El término *licencia* se refiere a las condiciones legales bajo las cuales se proporciona la obra.

- El término *dominio público* denota la ausencia de derechos de autor y restricciones similares, ya sea por omisión, ya por renuncia a todas esas condiciones.

El Grupo de Trabajo de Datos Abiertos en la Ciencia del OKF se creó en 2009 con la finalidad de desarrollar directrices, herramientas y aplicaciones para promover datos abiertos en la ciencia y permitir a los científicos maximizar el uso e impacto de ellos. En la actualidad, está integrado por una comunidad diversa de científicos, recopiladores de datos, abogados y otras personas interesadas tanto en los datos abiertos como en el concepto amplio de ciencia abierta (Molloy, 2011).

La definición de abierto que cristaliza en el OKD lleva aparejada la libertad para utilizar, reutilizar y redistribuir sin restricciones más allá de las exigencias de atribución y de compartir igual. Cualquier restricción adicional convierte el artículo en conocimiento cerrado. Enfatiza, además, la importancia de la usabilidad y el acceso a la totalidad del conjunto de datos o trabajo de conocimiento:

El trabajo deberá estar disponible en su conjunto y con un coste de reproducción razonable, pudiéndose descargar preferiblemente a través de Internet sin cargo alguno. El trabajo también deberá estar disponible en una forma adecuada y modificable.

Estos condicionantes son muy relevantes en lo que respecta a los datos científicos, puesto que, en algunos supuestos, puede accederse a los datos en suplementos en línea añadidos a las publicaciones, pero no se dispone de licencia para poderlos reutilizar o bien es accesible y puede ser reutilizado de una manera que no permita ni capturarlos ni modificarlos. Antes de que existieran los suplementos en línea, solicitar y obtener los permisos y los datos era una tarea ardua. No obstante, incluso hoy en día, con la posibilidad de descargar la información no resulta tan sencillo discernir qué derechos tiene cada uno a la hora de reutilizar. Es más, esta dilucidación puede resultar muy confusa, fruto de la inexistencia de licencias y de claridad en los términos de uso. En algunos casos, los datos en los que se apoya el artículo son públicos, incluso cuando no lo es el propio artículo, aunque en muchos supuestos no se haga explícitamente. Por eso, resulta de vital importancia etiquetar y licenciar claramente para evitar que los científicos dediquen horas a investigar si los datos son abiertos o no.

John Wilbanks, miembro de Creative Commons, y algunos miembros clave del OKF (Rufus Pollock y Peter Murray-Rust de la Universidad de Cambridge y Cameron Neylon) se dedicaron durante dos años al desarrollo de un conjunto de principios para publicar datos científicos en abierto, utilizando el OKF y el

protocolo de Science Commons para implementar datos de acceso abierto⁵² como directrices y precedentes (Science Common, s. f.).

En los supuestos en los que la financiación pública sirva para el desarrollo de la ciencia existe, cuando menos, una obligación de dar acceso a la información aparejada. Esta declaración tan evidente se desarrolla de forma mucho más elegante y concreta en los Principios de Panton, que establecen que la ciencia pública no debería tener licencias ni controles de innovación, y debería poder reutilizarse a nivel mundial (Clearwater, 2010).

Este es el origen de los Principios de Panton, que reciben el nombre del *pub* Panton Arms en Cambridge, donde se celebraron la mayor parte de las reuniones de trabajo para su elaboración (Molloy, 2011). Fueron publicados en febrero de 2010 y ya se han adherido a ellos más de ciento cincuenta miembros. Su propósito principal era hacer accesibles los datos científicos publicados (Ogburn, 2016).

El ámbito de estos principios cubre todos los datos primarios experimentales publicados en los escritos de investigación, incluidos los contenidos de datos de cualquier tabla o gráfico y de todas las imágenes, audios o videos que sirven de mecanismos primarios para la captura de datos.

Los Principios de Panton exigen que todos estos datos, con muy pocas excepciones, se pongan expresamente en el dominio público. Aunque entienden que existen algunas excepciones que pueden justificar que no se compartan los datos, por ejemplo: en caso de que ello pudiera vulnerar la privacidad del paciente, o bien que no se dé la localización de especies en peligro de extinción.

Autores de estos principios como Mayer (s. f.) se han adherido a varios principios con la finalidad de poner los datos a disposición de todos de forma abierta explicitando además los valores considerados y los usuarios a los que se dirige así como los usos:

- Los editores deben hacer una declaración explícita y sólida (renuncia, licencia) de sus deseos y expectativas con respecto a la reutilización y redefinición de los datos y sus subconjuntos.
- Muchas licencias no son apropiadas para datos o colecciones de datos, ejemplo de ello son las licencias Creative Commons, con la salvedad de la CC0. Se pide a los editores que utilicen modelos de licencias promovidos por la definición abierta, como PDDL, ODC-BY u ODbL.

⁵² Disponible en <http://sciencecommons.org/projects/publishing/open-access-data-protocol>. (Consultado por última vez el 15 de diciembre de 2016).

- Además, se desaconseja limitar la reutilización comercial, la reorientación o la exclusión de usos o usuarios particulares, debido a la posible preservación de los datos.
- Los datos de investigación financiados con fondos públicos deben ponerse explícitamente en el dominio público mediante el uso de la dedicación y de la licencia de dominio público o del uso del *wavier* CC0. Estos principios abordan explícitamente cuestiones técnicas, principalmente desde la perspectiva de la reutilización y la reorientación, así como del doble rol de los investigadores: ser generadores de datos de investigación y al mismo tiempo usuarios de los mismos.

4.3.3.4. The FAIR Guiding Data Principles

Sin lugar a dudas, existe una necesidad cada vez más urgente de mejorar las infraestructuras sobre las que se apoya la reutilización de datos científicos (Wilkinson *et al.*, 2016). Con esta finalidad, una serie grupos de interés relacionados con la ciencia abordaron la redacción de los Fair Guiding Data Principles (Force 11, s. f., The FAIR Data Principles).

Aunque una buena gestión de datos no es un fin en sí mismo, es la clave para conducir el descubrimiento del conocimiento y, por ende, de la obtención de los datos en cuestión y de su integración así como de su reutilización por parte de la comunidad (Wilkinson *et al.*, 2016).

El preámbulo de los Fair Data Principles establecía que uno de los grandes retos de la ciencia intensiva en datos era facilitar el descubrimiento del conocimiento ayudando tanto a las máquinas como a las personas al descubrimiento, acceso, integración y análisis de las tareas relacionadas con los datos científicos y sus algoritmos asociados y flujos de datos. Como se señaló más arriba, se describe FAIR como un conjunto de principios rectores con la finalidad de hacer los datos fáciles de encontrar (*findable*), accesibles (*accessible*), interoperables (*interoperable*) y reutilizables (*reusable*).

En el ecosistema de la e-ciencia (Force 11, s. f., Guiding Principles...), el reto de posibilitar el uso de los datos de investigación y los métodos representa una enorme complejidad. Existen una multiplicidad de *stakeholders*. Por una parte, se encuentran los investigadores que quieren compartir sus datos y sus interpretaciones y los editores profesionales de datos que ofrecen sus servicios, *software* y *tool-builders* que proporcionan servicios de análisis y procesamiento de datos.

Por otra parte, están los organismos de financiación (privados y públicos) cada vez más preocupados por la correcta administración de los datos y una comunidad de

minería de datos científicos que integra y analiza los *outputs* para acelerar los hallazgos. El análisis computacional para descubrir patrones significativos en grandes conjuntos de datos vinculados se está convirtiendo rápidamente en una actividad rutinaria de investigación. Proporcionar datos legibles máquina-máquina como el sustrato principal para el descubrimiento del conocimiento y para que estos procesos relacionados con la e-ciencia funcionen de forma fluida y sostenible es uno de los grandes retos de la e-ciencia.

En enero de 2014, a petición de Centro Holandés de e-ciencia y del Centro Tecnológico Holandés de Ciencias de la Vida (DTL), se reunieron en el centro Lorentz de Leiden (Holanda) representantes de todos los grupos con la finalidad de reflexionar y debatir de qué manera se podía seguir potenciando este ecosistema. De estos intercambios, surgió la idea de definir, apoyar y diseminar ampliamente un conjunto mínimo de principios y prácticas acordadas, para los proveedores de datos y los usuarios de los mismos —máquinas como personas— con la finalidad de encontrar más fácilmente, acceder, interoperar y reutilizar, citándolos de la forma adecuada, las enormes cantidades de información generadas en la actualidad.

Estos sencillos principios y prácticas tienen como finalidad permitir que se de una más amplia clase de comportamientos integrativos y exploratorios que soporten un amplio rango de opciones de tecnología y de implementaciones, de igual manera que el protocolo de Internet (IP) proporcionó una capa mínima que permitió la creación de una amplia gama de suministro de datos, el consumo y las herramientas de visualización en Internet.

Para que resulten *fáciles de encontrar (findable)*:

- F1. Se asigna a los metadatos de forma global y eterna un único identificador digital.
- F2. Se describen los datos con metadatos enriquecidos.
- F3. Los metadatos quedan registrados o indexados en fuentes que pueden consultarse.
- F4. Los metadatos especifican el identificador de los datos.

Para que resulten *accesibles (accessible)*:

- A1. Pueden recuperarse los metadatos usando su identificador persistente usando un protocolo de comunicaciones estandarizado.
 - A1.1. El protocolo es abierto, libre y de implementación universal.

A1.2. El protocolo permite un procedimiento de autenticación y autorización, en los casos en los que resulte necesario.

A2. Los metadatos son accesibles, incluso en el supuesto de que los datos ya no estén disponibles.

Para ser *interoperables* (*interoperable*):

I1. Los metadatos usan un lenguaje formal, accesible, compartido y de amplia aplicación para la representación del conocimiento.

I2. Los metadatos usan vocabulario que sigue los principios FAIR.

I3. Los metadatos incluyen referencias cualificadas a otros metadatos.

Para ser *reutilizables* (*reusable*):

R1. Los metadatos disponen de una pluralidad de atributos relevantes y precisos.

R1.1. Los metadatos se publican con una licencia de uso clara y accesible

R1.2. Los metadatos están asociados a su procedencia.

R1.3. Los metadatos cumplen con los estándares relevantes de este campo.

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

Figura 49.

Cómo aplicar los principios FAIR. Fuente: Wilkinson *et al.*, 2016.

La adopción de los principios se está extendiendo y muchas instituciones han optado por la aplicación de algunos de ellos. Incluso el Programa Horizonte 2020 los ha adoptado como base en sus políticas.

De igual manera, el 6 de septiembre de 2016, el G20 endosó los FAIR Principles. También, la ciencia en Europa ha acogido los FAIR Principles como base para compartir datos administrativos financiados (Mons *et al.*, 2017). Parece que los FAIR Principles están ampliamente adoptados; sin embargo, con el desarrollo de Internet, están surgiendo discrepancias sobre lo que son y para qué sirven. Hasta qué punto hay consenso es discutible.

Por eso, están surgiendo grupos de discusión que quieren ir aclarándolos, reinterpretándolos. En efecto, el paso de la ciencia a la nube y, por lo tanto, a que la investigación se haga de una manera más abierta y compartida, hace que los FAIR Principles sean todavía más necesarios (Mons *et al.*, 2017). No cabe duda de que meollo de estos principios está en la reutilización de los datos, de manera que sean lo más valiosos posibles. De hecho, no definen estándares técnicos para no establecer restricciones, sino todo lo contrario: impulsan que se reutilicen lo máximo posible.

La accesibilidad de FAIR es diferente de *open* (Mons *et al.*, 2017); es accesibilidad en virtud de condiciones bien definidas. Por lo tanto, existen razones que pueden legitimar que se protejan datos y servicios obtenidos con financiación pública del acceso público.

Los niveles de reutilización todavía no son muy altos y por ello se ha creado el término *reuseless* (inútiles para su reutilización) para referirse a estos datos.

El FAIR sigue planteando algún problema. En efecto, está pensado para que el que tiene el dato lo haga *open access*. Una vez que el que tiene el dato los hace «reusable» ¿los mantiene?, ¿los actualiza?, ¿los pone al día? El que tiene el dato no impide el uso, pero, si cierra el instituto o la página web, esos datos dejan de ser reusables. Incluso en el mejor escenario —que es que el que tiene el dato lo abra lo máximo posible— necesita el apoyo de estructuras públicas que mantengan los datos accesibles. Los FAIR están pensados desde un punto de vista subjetivo, y desde ese punto de vista está muy bien, pero no es suficiente. No se puede obligar a un investigador a tener una infraestructura de soporte a los datos.

El FAIR está muy bien pensado como cambio cultural y como principios a los generadores de datos. Para que la política FAIR funcione son necesarias las *data management policies*. Es lo que deberían hacer todas las instituciones. Un ejemplo de ello puede encontrarse en la Open Science Cloud europea (EOSC).

La obligación del generador de los datos es entregarlos en FAIR. Para ello se crean lo que se llaman los *branding*, que permiten reconocer rápidamente el nivel de FAIR que se ha aplicado a los datos.

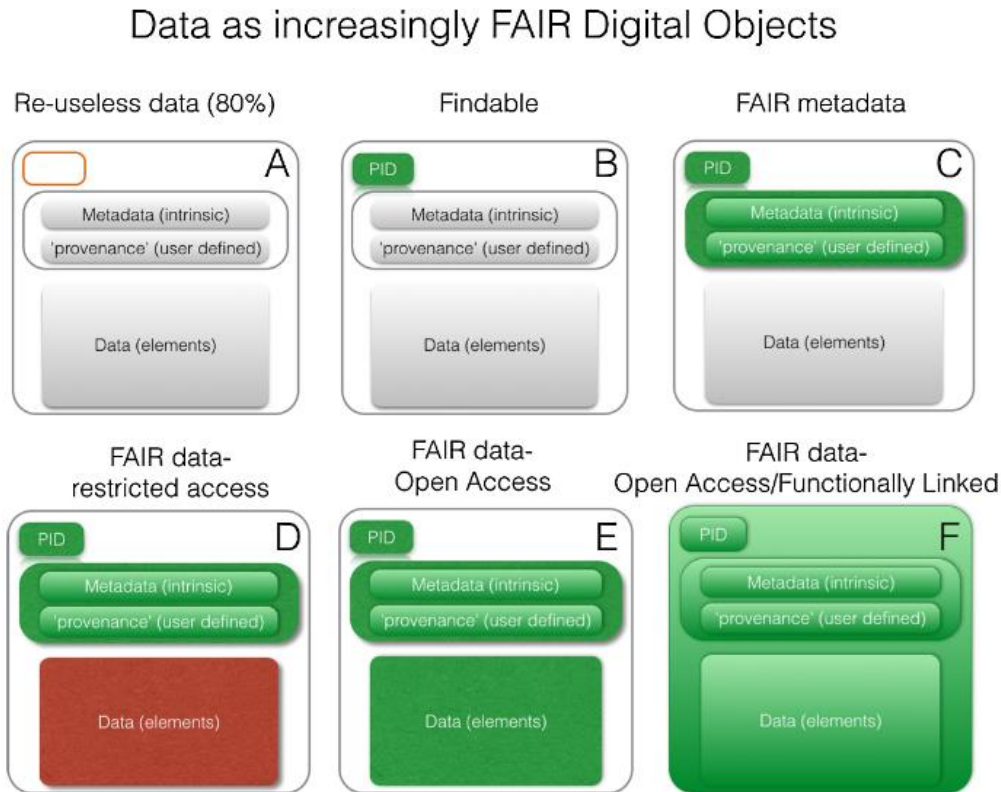


Figura 50. Distintos niveles de *FAIRness*. Cuantos más elementos estén coloreados más *FAIR* son. Por ejemplo, si se añade un identificador persistente mejora el nivel de *FAIRness* de ese elemento. Los elementos en verde representan elementos FAIR y *open*, los elementos en rojo FAIR y *closed*

Aunque los principios describen lo que es deseable, dan pocas directrices en relación con la forma de llevarlo a cabo y esto es problemático en este momento en el que grandes organizaciones se están adhiriendo a los mismos (Wilkinson *et al.*, 2016).

Puede que se hayan asentado las bases, pero queda pendiente el aterrizar estos principios a la realidad actual e ir obteniendo los compromisos de los distintos *stakeholders*. El camino acaba de empezar.

4.4. Dominio público: ¿alternativa al *copyleft* o a las licencias Creative Commons? ¿Primeros pasos hacia la interoperabilidad global?

¿Qué debe ser público y qué debe ser privado en la investigación científica? (Eisemberg y Nelson, 2002). ¿Cómo avanzará más la ciencia, poniendo el conocimiento en el dominio público o protegiéndolo con los derechos de propiedad intelectual o *copyright*?

Todas estas cuestiones y muchas más que se han ido planteando a lo largo de este trabajo llevan a analizar cómo el poner los datos y los descubrimientos en el dominio público puede ser uno de los primeros pasos hacia la interoperabilidad legal.

En efecto, lo que ha quedado claro es la complejidad de regímenes y restricciones aplicables a los datos. Esta confluencia de regímenes genera una confusión en los científicos que hace que, bien no reutilicen los datos (por desconocimiento), bien los usen sin respetar las normas que los protegen (por esa misma razón).

Como se ha puesto de manifiesto a lo largo de todo este trabajo, los datos están ahora en el mercado. No compartirlos hoy parece tarea imposible y poco productiva. Por esa razón, se habla de *open science*.

La cuestión es si procede hacer una defensa del acceso público a los datos como parte de la preservación de la ciencia. Dreyfuss (2004) no es el único autor que señala que tal vez el sistema de patentes y protección no sea ya el adecuado en el marco de la e-ciencia. En efecto, señala que en la actualidad existe un mar de protección cerrada salpicado por islas de dominio público.

No obstante, como se ha puesto de manifiesto en el punto anterior relacionado con las distintas propuestas para ir construyendo una interoperabilidad global, existen múltiples organizaciones que promueven principios tendentes a poner los datos en el dominio público y hacerlos accesibles.

La mayor parte de las personas están de acuerdo en que el dominio público y el acceso abierto a los datos científicos son clave hoy para acceder al conocimiento (Foray, 2004). El primer paso para avanzar hacia modelos de innovación es reconocer el dominio público que representa espacios públicos que conducen a la eficiente difusión del conocimiento. De hecho, las acciones colectivas son importantes a la hora de abordar el concepto de dominio público.

El dominio público es un régimen históricamente más presente en Estados Unidos, que es donde es más amplio como consecuencia de, por ejemplo, las políticas de *public access*. Sin embargo, este concepto se está extendiendo a más lugares y, así, por ejemplo la Unión Europea ya lo está adoptando paulatinamente. Como se ha puesto de manifiesto en los puntos anteriores, tanto los programas europeos de financiación de proyectos científicos en el marco del Horizonte 2020 como los estadounidenses en el marco de las políticas de la National Science Foundation y del National Institute of Health exigen poner los datos en abierto, con la finalidad de que los descubrimientos realizados sean accesibles a todos.

De alguna manera, el recurrir al dominio público simplifica las tareas de los investigadores que pueden reutilizar los datos sin restricciones, con lo que se consigue una interoperabilidad legal total.

Aunque el concepto de dominio público nace en los Estados Unidos con la US Copyright Act que especifica que la protección del *copyright* tiene una duración limitada antes de que el transcurso del tiempo la ponga en el dominio público. Una vez que el trabajo esté en el dominio público podrá usarse sin permiso y sin necesidad de pagar ninguna cantidad. El razonamiento que existe detrás es que la normativa del *copyright* tiene que establecer un equilibrio entre las necesidades del creador del trabajo — obtener un beneficio— y las necesidades de la sociedad — que se desarrolle el conocimiento humano mediante el uso y adaptación de los trabajos de otros—.

Las normas por las que se rige el dominio público han ido cambiando. La protección al amparo del *copyright* es cada vez más larga y ahora se protege incluso en el supuesto de que el creador no haya hecho ningún esfuerzo por registrar su trabajo, pues la inscripción en un registro ha dejado de ser un requisito formar constitutivo.

Como consecuencia de todos estos cambios, resulta complicado saber si una obra está en el dominio público. Como ya se ha señalado en varias ocasiones a lo largo de este trabajo, la multiplicación de los datos ha supuesto, sin duda, un cambio en el ámbito de la investigación científica.

Si bien han surgido un gran número de tipos de licencias *free*, *open* o libres, las licencias suponen el reconocimiento, presuponen la existencia de derechos de propiedad intelectual o *copyright*. Wilbanks (2008) señala que existe un conflicto entre la simplicidad y la seguridad jurídica, y propone para afrontarlo la adopción de dos medidas: la reconstrucción del dominio público y el uso de normas científicas en las que se exprese el deseo del proveedor de datos.

La reconstrucción del dominio público puede conseguirse mediante el uso de una herramienta jurídica según la cual los conjuntos de datos de investigación se encuentran en el dominio público si se cumple cualquiera de las siguientes condiciones (RDA-CODATA Legal Interoperability Interest Group, 2016):

1. El tiempo de protección de propiedad intelectual, en caso de ser de aplicación, ya haya transcurrido (por ejemplo, de conformidad con el acuerdo internacional en *copyright law*, es la vida del autor, además de un mínimo de cincuenta años). En la legislación nacional de muchos países (incluidos todos los países de la OCDE, salvo Canadá), el término se ha extendido a la vida del autor, más setenta años. Existe una norma especial en Estados Unidos en el supuesto de *works for hire* (es decir, de los trabajos

que realicen las personas contratadas en el marco de su contrato), según la cual el plazo será de noventa y cinco años desde la fecha de la primera publicación o ciento veinte años desde la fecha de su creación (primará la que termine antes). De conformidad con las regulaciones de protección de bases de datos, el plazo original de protección de propiedad exclusiva es de quince años, que se puede renovar quince años más con cualquier inversión sustancial en el mantenimiento de la base de datos.

2. No puede protegerse el objeto para empezar. En la *Copyright Law*, por ejemplo, no son susceptibles de ser protegidos los hechos de la naturaleza, que representan una gran parte del contenido de los conjuntos de datos.
3. En los supuestos en los que el titular de los datos renuncie a todos los derechos (*copyright* y otros derechos de propiedad intelectual, pero no a los derechos morales de autor) que tenga sobre los datos para siempre.
4. Cuando un Gobierno, por el tipo de información, la excluya de los derechos de *copyright* o de propiedad intelectual, sea mediante un tratado, un acuerdo, una política o regulaciones al respecto.

Una de las cuestiones fundamentales que se plantea en relación con el dominio público es la mercantilización de los datos (Reichman y Uhler, 1999), que da lugar a la creación de prometedores nuevos mercados para mercancías relacionadas con la información y abre nuevas posibilidades para que se usen los datos y la información con fines científicos y educativos.

Además, lo que está ocurriendo hoy es que se está desarrollando una cultura de compartir construida en el hecho de no reclamar las infracciones contra el *copyright* incluso en el supuesto de que existan.

4.5. ¿Centralización de los datos?

En el mundo de hoy, parece que están prosperando los diseños *by design*. Ya se habla de *privacy by design*. De alguna forma, este concepto responsabiliza al titular de los datos del uso que haga de los mismos, y exige una preconfiguración de privacidad. Algunos autores como Contreras y Reichman (2015) extienden este concepto al intercambio de datos y hablan de *sharing by design*.

De alguna manera, se está trabajando a todos los niveles en la integración de los datos. Al igual que en el análisis sobre si tiene sentido que existan tantas organizaciones recabando datos de biodiversidad, se plantea en este ámbito la multiplicidad de herramientas y políticas, existentes a la hora de recabar y compartir datos. La simplificación es lo que persigue, de alguna manera, la interoperabilidad legal. En lugar de que los datos se almacenen en lugares distintos, la idea es poder

acceder a ellos con diversas herramientas y a través de ciertas fuentes y que ambas sean contrastables (Contreras y Reichman, 2015).

Structural models for scientific data pools				
Data-sharing options				
BENEFITS AND COSTS	CENTRALIZED	INTERMEDIATE DISTRIBUTED	FULLY DISTRIBUTED	NONCOMMONS
Incremental research benefits				
Data access	Access to all data in unified manner	Access to multiple repositories through central portal	Access to each repository separately, but under a common usage/access policy and single approval	Ad hoc coordination with other repositories only
Data analytics	Most powerful search, analysis, quality assurance of aggregated data	Cross-repository searching and analytics; Metadata and aggregate statistics can be developed by central authority	Index/catalog only	Index/catalog only
Costs				
Up-front costs	Structure and build centralized repository; Develop data interoperability mechanisms; Develop common usage policy	Develop data interoperability mechanisms; Develop common usage policy	Develop common usage policy	Few up-front costs
Ongoing centralized costs	Operating and maintaining central repository; administering policies	Operating and maintaining portal; administering policies	Administering policies	No central costs
Ongoing distributed costs	Few distributed costs	Operating and maintaining repositories	Operating and maintaining repositories	Operating and maintaining repositories
Governance overhead	Central repository	Central portal/services, each distributed repository, and interrelationships	Each distributed repository and interrelationships	Each distributed repository with minimal coordination
Structural models for scientific data poolsData-sharing options				

Figura 51. Conjuntos estructurales de conjuntos de datos científicos. Fuente: Contreras y Reichman, 2015.

Capítulo 2. Segundo marco de contextualización: la diversidad biológica, también llamada biodiversidad

Dentro de este marco de contextualización, se analizará el concepto de diversidad biológica o biodiversidad. Se hará en dos pasos: primero, con un análisis del concepto teórico y, seguidamente, con el análisis de la interacción de distintos instrumentos, plataformas, organismos, etc.

Los primeros pasos hacia la concienciación sobre la diversidad biológica tuvieron lugar en los años setenta. Con los años han ido multiplicándose las iniciativas para gestionar las cuestiones relacionadas con la diversidad biológica en distintos ámbitos. En este punto es importante entender de qué manera la creación, la evolución y la falta de consecución de los objetivos en algunos casos, incluso se habla de fracaso, han servido de trampolín a nuevas iniciativas y de qué manera estas están conectadas.

Una de las grandes cuestiones que surgió en este marco es la definición del objetivo perseguido y la evaluación para saber si se ha alcanzado.

1. LA CONCIENCIACIÓN EN MATERIA DE BIODIVERSIDAD: ORÍGENES DEL TÉRMINO

El término «biodiversidad» es un calco del inglés *biodiversity* que, a su vez, resulta de la contracción de la expresión *biological diversity*. Apareció en los años ochenta en la preparación de una reunión auspiciada por la National Science Foundation en Washington: el National Forum on BioDiversity. Fue acuñado Walter G. Rosen, quien considera que el término «biodiversidad» resulta más usable y sencillo que «diversidad biológica». Este concepto de biodiversidad que citan Harper y Hawksworth (1995) lo recogió también Wilson (1988).

Más allá del propio término, en los años setenta empezó a despertarse una preocupación por la conservación de la diversidad biológica. Comenzaron a aparecer ya iniciativas intergubernamentales, por ejemplo, la Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres (CITES) firmada en 1973 en Washington.

En la literatura científica anterior a los ochenta, la diversidad biológica aludía a la diversidad de especies en tanto que característica estructural de los ecosistemas (Toledo, 1994). La explosión significativa del concepto de biodiversidad emanó básicamente de dos publicaciones aparecidas en 1980 (Nuñez *et al.*, 2003). De hecho, en un informe al presidente de los Estados Unidos, Lovejoy (1980) lo definía sin nombrar el término. El otro, prosiguen Norse y McManus (1980) quienes colaboraran en el Consejo en Calidad Ambiental de la Casa Blanca (Nuñez *et al.*, 2003). Además, durante ese mandato elaboraron, también, un capítulo para el 11.º Reporte Anual del Consejo en Calidad Ambiental (Jeffries, 1997; Harper y Hawksworth, 1995) en el que se examina la biodiversidad global, definiéndola con dos conceptos relacionados entre sí: diversidad genética (la cantidad de variabilidad genética dentro de las especies) y diversidad ecológica (el número de especies en una comunidad de organismos).

Otra definición, al examinar el concepto de biodiversidad global, incluyó dos conceptos relacionados entre sí: diversidad genética —la cantidad de variabilidad genética dentro de las especies— y diversidad ecológica —el número de especies en una comunidad de organismos— (Norse y McManus, 1980). En efecto, la biodiversidad no era solo objeto de estudio por parte de la biología (Jeffries, 1997).

El panorama actual mostraba una creciente degradación y agotamiento de los sistemas biológicos y de su diversidad (Nuñez *et al.*, 2003). El término diversidad biológica nació indisolublemente ligado a las instituciones académicas y organismos nacionales e internacionales dedicados a la conservación biológica y como un concepto sintético que incluía por igual enfoques de taxonomía, ecología y biogeografía (Toledo, 1994).

De esta manera, quedó superada la asociación de la biodiversidad con la diversidad de especies, pues se recogía un concepto mucho más amplio. La biodiversidad, por consiguiente, resultó de una integración de la historia natural —o el conocimiento de qué especies existían—, de la medida —o de qué manera se evaluaba lo que se quería decir con diversidad y cómo se comparaban las zonas— y el uso —o el darse cuenta de la importancia de algunas de las especies para la supervivencia del *Homo sapiens*— (Usher, 1991). Llegó a decirse que biodiversidad era el sinónimo científico de naturaleza (Takacs, 1996).

Dos razones fundamentales justificaron que la diversidad biológica o biodiversidad suscitase tal nivel de interés en los años ochenta, entre los científicos y entre el público en general. La primera, por supuesto, la preparación y celebración del National Forum on Biodiversity en Washington y, la segunda, la creciente toma de conciencia de la estrecha relación existente entre la conservación de la biodiversidad y el desarrollo económico (Hannigan, 2014).

Teniendo en cuenta esta idea, convergieron fundamentalmente en los países del sur dos conceptos: desarrollo y preservación de los ecosistemas, dado que en esos países la aceleración de pérdidas de ecosistemas planteó importantes problemas de carácter sociopolítico.

La diversidad biológica, entendida como la variabilidad de seres vivos y de los ecosistemas en los que viven, ha sido el pilar sobre el que se han construido las civilizaciones humanas. Su conservación ha sido la base del desarrollo sostenible y, como tal, es uno de los grandes desafíos de la era moderna.

Con esos condicionantes, una serie de biólogos y científicos comenzaron a defender que la pérdida de la biodiversidad era un problema global y se dieron cuenta de la necesidad de hacer una aproximación a las distintas cuestiones de forma más global. Empezaron a agruparse biólogos con inquietudes comunes que identificaban cuestiones pendientes de resolver. El término «biodiversidad» quedó, por lo tanto, perfectamente integrado en el pensamiento conservacionista (Wilson, 1988). El concepto de biodiversidad y la preocupación por su pérdida son muy recientes, aunque los naturalistas lleven más de doscientos cincuenta años estudiándolo (Halffter, 1994).

En la actualidad, los paisajes están empezando a considerarse incluidos también en el concepto de biodiversidad, después de muchos años de discrepancias en los que muchos biólogos defendían que los paisajes eran para los pintores. Sin embargo, ya los científicos empiezan a reconocer el paisaje como una escala superior de biodiversidad que no deja de ser la escala que agrupa los ecosistemas. Se pueden tener ecosistemas parecidos y que los paisajes sean diferentes. Cambia la estructura de donde está uno y donde está otro. En realidad, el paisaje influye en los ecosistemas (Tucker, *et al.*, 2018). Por lo tanto, la conservación de la biodiversidad requiere también la protección de los paisajes ecológicos. En España, de hecho, la Ley 42/2007, de 13 de diciembre, del Patrimonio Natural y de la Biodiversidad ya incluye las referencias de los paisajes.

En paralelo, la proliferación de datos hizo aparecer la necesidad de ponerlos en común, motivo por el cual se crearon diferentes instrumentos. Surgió, por consiguiente, la necesidad de recabar la información que permitiera tomar decisiones informadas y parar o reducir la pérdida de la biodiversidad y de los ecosistemas. Para lograrlo, es importante entender cómo se configuran y de qué manera interactúan.

2. EL CONVENIO SOBRE LA DIVERSIDAD BIOLÓGICA (CDB)

2.1. La gestación del CDB

Como consecuencia de todos estos movimientos de concienciación de la biodiversidad, a finales de los años ochenta empezaron a fraguarse en paralelo dos grupos: uno de cambio climático (que daría lugar a la Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC) fue adoptada en Nueva York el 9 de mayo de 1992 y entró en vigor el 21 de marzo de 1994. Permite, entre otras cosas, reforzar la conciencia pública, a escala mundial, de los problemas relacionados con el cambio climático) y otro de biodiversidad (lo que finalmente sería el CDB) (Naciones Unidas, 1992).

La relación entre el contenido científico y el compromiso político es clave en lo que respecta a las cuestiones de medio ambiente (Miller, 2001), pero rara vez los políticos resultan elegidos por proteger el medio ambiente (Watson, 2001).

Tal vez esto explique, de alguna manera, la distancia que separa lo político de lo científico. Naciones Unidas a través del PNUMA (Programa de las Naciones Unidas para el Medioambiente) creó un grupo de trabajo para evaluar la oportunidad y viabilidad de un convenio paraguas que le permitiera racionalizar las actividades actuales en el ámbito de los convenios internacionales del hábitat para la flora y fauna silvestre (Unión Mundial para la Naturaleza, 1963) y abordar otras áreas que pudiesen quedar enmarcadas dentro de este convenio (Tolba, 1998).

El CDB fue un paso importante y de gran trascendencia; de hecho, constituyó un compromiso histórico de muchos países a nivel mundial (Gaston y Spicer, 1998). A día de hoy lo han firmado más de ciento noventa países, aunque no todos lo han ratificado. Recogió concesiones de los países desarrollados que nunca se habían logrado antes en ninguna otra negociación multilateral (Le Prestre, 2002a; Désirée, 2002).

Los abogados internacionalistas lo incluyeron dentro de una nueva generación de instrumentos legales cuya finalidad era reconciliar los desarrollos imperativos del Sur con las exigencias del Norte (Tinker, 1995). En realidad, el ámbito de aplicación del CDB va mucho más allá de la conservación y contempla el desarrollo sostenible y las cuestiones de equidad (LePrestre, 2008).

La biodiversidad, en efecto, se ha convertido en una cuestión de carácter global. Hay que aprender a construir un bien colectivo, a compartir una visión de índole planetaria de los fenómenos económicos y ecológicos (Aubertin *et al.*, 1998).

El origen de las negociaciones se remonta a 1987, con la Decisión 14/86 (Naciones Unidas, s. f.) del Consejo de Administración del PNUMA que decidió constituir un

grupo de expertos en biodiversidad a efectos de armonizar todos los convenios existentes relacionados con el tema (Secretariat of the Convention on Biological Diversity, 2005). En la primera reunión de este grupo de expertos, quedó clara la necesidad de crear un instrumento internacional vinculante para la conservación y uso sostenible de la diversidad biológica.

En 1989 se estableció otro grupo *ad hoc* con esta misma idea: crear un instrumento internacional vinculante para la conservación y el uso sostenible de la diversidad biológica, pero se llegó un paso más allá. Se planteó la necesidad de hacer un reparto de costes y beneficios entre países desarrollados y países en vías de desarrollo, y los medios y maneras de proteger la innovación de las comunidades locales. Este grupo de trabajo creado *ad hoc* se convirtió, en 1991, en el comité intergubernamental de negociación y, después de siete reuniones de trabajo de las cuales cinco fueron de negociación, sus integrantes llegaron a un acuerdo sobre el texto, aprobado en el acta final de la Conferencia de Nairobi. El convenio quedó abierto para su firma en la Conferencia de las Naciones Unidas sobre el Medio Ambiente y el Desarrollo (CNUMAD), en Río de Janeiro en 1992, y entró en vigor en 1993.

Nació como un tratado internacional jurídicamente vinculante. Los objetivos fundamentales que estableció fueron: la conservación de la biodiversidad, el uso sostenible de sus componentes y la participación justa y equitativa en los beneficios derivados de la utilización de los recursos genéticos (Naciones Unidas, s. f., Convenio sobre la Diversidad Biológica).

En última instancia, quería promover medidas conducentes a un futuro sostenible. Para ello, era necesario elaborar indicadores de biodiversidad y de los servicios de los ecosistemas (Balmford y Bennun, 2005). Señalaron, incluso, la importancia de que los conservacionistas aprendieran de los economistas que han conseguido elaborar indicadores más claros en relación con el funcionamiento del mercado.

El artículo 2 del CDB definió la «diversidad biológica» como la variabilidad de organismos vivos de cualquier fuente, incluidos, entre otras cosas, los ecosistemas terrestres y marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte; comprende, así, la diversidad dentro de cada especie, entre las especies y de los ecosistemas.

Toda una serie de características convirtieron el texto del CDB en un documento⁵³ simultáneamente legal y político: la amplitud que abarca, la complejidad y el

⁵³ Disponible en

http://graduateinstitute.ch/files/live/sites/iheid/files/sites/developpement/shared/developpement/mdev/soutienauxcours0809/hufty_biodiversite/8_McGraw-LePrestre.pdf. (Consultado por última vez el 22 de enero de 2015).

compromiso (Désirée, 2002). Se configuró como un acuerdo marco desde tres puntos de vista (Le Prestre, 2002a).

En primer lugar, creó una estructura internacional para promover la cooperación internacional y para contribuir a su implementación en los Estados. Estableció, por lo tanto, un marco general y flexible de obligaciones que las partes deberán desarrollar en sus leyes nacionales y en sus políticas.

En segundo lugar, permitió realizar subsiguientes desarrollos del mismo mediante la negociación de anexos y protocolos. Este instrumento se usó con frecuencia para crear unos marcos generales de *soft law* que permitiera, en posteriores negociaciones, incluir requerimientos más concretos. Por último, construyó o completó acuerdos ya existentes.

El principal problema planteado radicaba en que la riqueza, en términos de biodiversidad, se encuentra en los países del Sur. Sin embargo, estos países no disponían de los medios necesarios para aprovechar estos recursos y, por lo tanto, eran los países desarrollados los que estaban haciendo estudios y aprovechando los recursos de los países del Sur.

Los países del Norte querían evitar que el desarrollo de los países del Sur destruyese la biodiversidad mientras que los últimos acusaban a los primeros de usar la justificación del medioambiente para frenar su desarrollo (Aubertin *et al.*, 1998).

Sobre ese punto pivotaba todo el concepto de participación justa y equitativa en los beneficios; tenía la finalidad de preservar esa riqueza y ese conocimiento de los países del Sur sobre la biodiversidad y evitar que los países más desarrollados pudiesen «apropiarse del conocimiento» y pasar a enriquecerse con ello haciendo que los países del Sur dejaran de obtener beneficios. Quedó regulado en el artículo 15 relativo al acceso a los recursos genéticos y al reparto de beneficios derivados de su utilización.⁵⁴ No es objeto de esta tesis estudiar la problemática de este artículo,

⁵⁴ BOE (1 de febrero de 1994). BOE-A-1994-2193, BOE, 27, 3113-3125: «Artículo 15. Acceso a los recursos genéticos.

1. En reconocimiento de los derechos soberanos de los Estados sobre sus recursos naturales, la facultad de regular el acceso a los recursos genéticos incumbe a los Gobiernos nacionales y está sometida a la legislación nacional.

2. Cada Parte Contratante procurará crear condiciones para facilitar a otras Partes Contratantes el acceso a los recursos genéticos para utilizaciones ambientalmente adecuadas, y no imponer restricciones contrarias a los objetivos del presente Convenio.

3. A los efectos del presente Convenio, los recursos genéticos suministrados por una Parte Contratante a los que se refieren este artículo y los artículos 16 y 19 son únicamente los suministrados por Partes Contratantes que son países de origen de esos recursos o por las Partes que hayan adquirido los recursos genéticos de conformidad con el presente Convenio.

4. Cuando se conceda acceso, este será en condiciones mutuamente convenidas y estará sometido a lo dispuesto en el presente artículo.

sino todo lo que queda excluido del mismo, aunque no podía dejar de mencionarse, tanto por la relevancia que tuvo y tiene como por la complejidad que presenta, tanta que podría ser objeto de una tesis separada.

Así, la erosión de la biodiversidad quedó incluida entre las cuestiones globales de medioambiente, lo que dejaba patente la necesidad de que existiera un compromiso generalizado. Aunque más de ciento ochenta Estados han ratificado el CDB, existe una falta de ratificación clave, merecedora de un desarrollo especial. En efecto, después de estar involucrado en toda la génesis del convenio y de firmar en primera instancia, Estados Unidos decidió, y sigue decidiendo hoy, no ratificar el CDB.

2.2. Estados Unidos: su negativa a ratificar el CDB

Dicha negativa merece una consideración específica por el significado que tiene, tanto externa como internamente. Dado que Estados Unidos fue un temprano impulsor de normas en materia de biodiversidad, por ejemplo la Ley de Especies en Peligro en 1973 (Estados Unidos, 1973), este comportamiento puede parecer inconsistente y no alineado con la política estadounidense general. En realidad, la respuesta de Estados Unidos no tiene una justificación única, sino que tuvo su origen en un complejo entramado de cuestiones (Blomquist, 2002).

No cabe duda de que Estados Unidos es un país convencido de la preservación de la biodiversidad biológica; aunque muestra una preocupación por todos los temas ligados al índice de pérdida de biodiversidad durante los años 1989-1990, en los años 1991-1992 no acaba de estar convencido sobre la conveniencia de firmar un tratado.

Sin embargo, tanto Bush, presidente republicano de Estados Unidos en el momento de la ratificación, como Clinton, demócrata y presidente posterior, mostraron su desacuerdo con algunas de las cuestiones clave adoptadas en el CDB, razón por la cual decidieron, en distintos momentos históricos, no firmar.

5. El acceso a los recursos genéticos estará sometido al consentimiento fundamentado previo de la Parte Contratante que proporciona los recursos a menos que esa Parte decida otra cosa.

6. Cada Parte Contratante procurará promover y realizar investigaciones científicas basadas en los recursos genéticos proporcionados por otras Partes Contratantes con la plena participación de esas Partes Contratantes y, de ser posible, en ellas.

7. Cada Parte Contratante tomará medidas legislativas, administrativas o de política, según proceda, de conformidad con los artículos 16 y 19 y, cuando sea necesario, por conducto del mecanismo financiero previsto en los artículos 20 y 21, para compartir en forma justa y equitativa los resultados de las actividades de investigación y desarrollo y los beneficios derivados de la utilización comercial y de otra índole de los recursos genéticos con la Parte Contratante que aporta esos recursos. Esa participación se llevará a cabo en condiciones mutuamente acordadas». Recuperado de <https://www.boe.es/buscar/doc.php?id=BOE-A-1994-2193>. (Consultado por última vez el 13 de julio de 2016).

La Administración de Bush, primero, decidió no ratificar el CDB esgrimiendo la necesidad de preservar las oportunidades de desarrollo económico de la industria estadounidense. Durante las discusiones habidas sobre la conveniencia o no de que Estados Unidos ratificase el CDB, surgieron toda una serie de cuestiones (Blomquist, 2002).

Después de realizar un profundo análisis de las distintas razones por las cuales Estados Unidos se negó a ratificar el CDB, Blomquist (2002) las estructuró, en cuatro fundamentales.

1. La *existencia* de una *tensión institucional* entre el *presidente* y el *Congreso* en relación con los asuntos extranjeros. El presidente de Estados Unidos tiene, de conformidad con la Constitución, la potestad de firmar tratados (y por ende de no firmarlos).

Al amparo de esta potestad (fuente de grandes discusiones institucionales), Bush decidió no ratificar el CDB, preocupado por el posible impacto que pudiera tener el CDB en su derecho constitucional. En efecto, de conformidad con este, la propiedad privada no es susceptible de apropiación pública sin compensación; por lo tanto, tuvo dudas de que pudieran hacerse concesiones relacionadas con la propiedad privada, por ejemplo: derechos de propiedad industrial de las compañías americanas.

2. La existencia de una *preocupación* de los conservadores en relación con la configuración emergente del *derecho medioambiental internacional*. Desde 1972, los tratados multilaterales se han ido multiplicando exponencialmente.

En Estados Unidos se ha ido desarrollando una corriente muy crítica en relación con los principios que subyacen en los tratados internacionales, por ejemplo, la Convención Marco de las Naciones Unidas sobre el Cambio Climático (Naciones Unidas, 1992) y el Protocolo de Kioto (Naciones Unidas, 2005); de hecho la Administración Trump, cuestionando la aplicación de estos acuerdos, ha decidido en 2017 salirse del Acuerdo de París de cambio climático.

Los estadounidenses eran, y son, más partidarios de una visión de la equidad medioambiental internacional alineada con los derechos nacionales soberanos de explotar recursos dentro de la jurisdicción o control de un país en relación con los derechos a los bienes comunes o compartidos (con independencia de que se trate de recursos naturales o de derechos de emisión) por orden de llegada. Esta forma de pensar entraba en conflicto con las nuevas premisas del nuevo orden internacional medioambiental que servía de base para el CDB.

3. Los *intereses* de las compañías estadounidenses de *maximizar los beneficios en biotecnología*. En efecto, la renuencia del Senado estadounidense a la ratificación del CDB tuvo mucho que ver con la relevancia política que tiene allí promover los intereses de las compañías estadounidenses en la obtención de lo que consideran beneficios legítimos en biotecnología mediante el desarrollo y comercialización de tecnologías útiles en lo que respecta a la agricultura, la industria farmacéutica y la medicina. El hecho de que Estados Unidos no ratificase el CDB pone en peligro el reparto de beneficios a nivel mundial (Vogel, 2005).

Aunque algunas de las preocupaciones que surgieron en la Administración de Bush en relación con la transferencia de tecnología y biotecnología se suavizaron bajo la dirección de la Administración Clinton, no desaparecieron. En efecto, los estadounidenses consideran que el texto final del CDB es poco claro, impreciso e inconsistente. En su opinión, las disposiciones del CDB abrieron la puerta a que normas internacionales pudieran entrar a regular la biotecnología estadounidense incluso en supuestos en los que la investigación, el desarrollo y el uso de la biotecnología fueran de ámbito exclusivamente doméstico.

Existió, también, una fuerte resistencia de una parte de la industria estadounidense, sobre todo de la farmacéutica. Adicionalmente, Bush tuvo en cuenta, también, el posible impacto del CDB en las obligaciones legales domésticas de materia de protección de la biodiversidad.

4. La *complejidad* existente para resolver las *repercusiones internacionales físicas y económicas*. Teniendo en cuenta las dificultades que existen dentro de las naciones soberanas como Estados Unidos a la hora de elaborar regulaciones eficientes que vayan más allá de los límites fronterizos de los Estados, no es sorprendente que lo que propone el CDB haya generado controversia, en general, y dentro de Estados Unidos, en particular.

En resumen, Estados Unidos no tenía disposición de ceder parte de su soberanía a la decisión de otros Estados ni de comprometer las oportunidades de desarrollo económico de la industria estadounidense.

A pesar de la variedad de razones esgrimidas, muchos autores pusieron de manifiesto que la razón fundamental por la que Estados Unidos se negó a firmar el CDB fue el tratamiento de los derechos de propiedad intelectual en la medida en que el CDB obligaba, fundamentalmente, a los países desarrollados (entre los que se encontraba Estados Unidos) a realizar una transferencia de tecnología a los países en vías de desarrollo.

De conformidad con el artículo 16 (c) del Tratado, Estados Unidos no solo tendría la obligación de transferir los productos tecnológicos comercialmente disponibles, sino también la propia tecnología, sin tener en cuenta los derechos de propiedad intelectual. Por lo tanto, la firma del CDB podía llegar a tener unos efectos muy perjudiciales para la industria estadounidense que esta no estaba dispuesta a asumir.

Incluso después de todo este análisis, sigue resultando complicado entender esa resistencia a firmar. Por una parte, Estados Unidos tenía muchas regulaciones internas con exigencias muy superiores a las exigidas en el propio CDB. Pero es que, además, no haber ratificado el CDB tuvo, y sigue teniendo hoy un coste para ellos: solo pueden asistir a las reuniones como oyentes y no pueden formar parte de la toma de decisiones (Kormos *et al.*, 2001).

De alguna forma, con la firma del tratado, Estados Unidos sentía que perdía, también, el control de los fondos dados a terceros Estados para la conservación de la biodiversidad, que estaba cediendo parte de su soberanía al efecto. Otra posible razón podía ser su cultura de política «aislada». El CDB no ha sido el único instrumento que Estados Unidos no ha ratificado: se ha negado a ratificar otros muchos de los tratados internacionales (Brunée, 2004), a pesar de su profunda involucración en temas medioambientales. Después de este paréntesis, procede realizar un análisis del convenio, comenzando por su estructura.

2.3. Los órganos de gobierno del CDB

La estructura de gobierno del CDB es la siguiente (Convention on Biological Diversity, s. f., Role). Aunque vaya a hacerse un análisis en detalle de cada una de las partes, resulta de utilidad ver la estructura representada gráficamente para comprender mejor el funcionamiento del CDB.

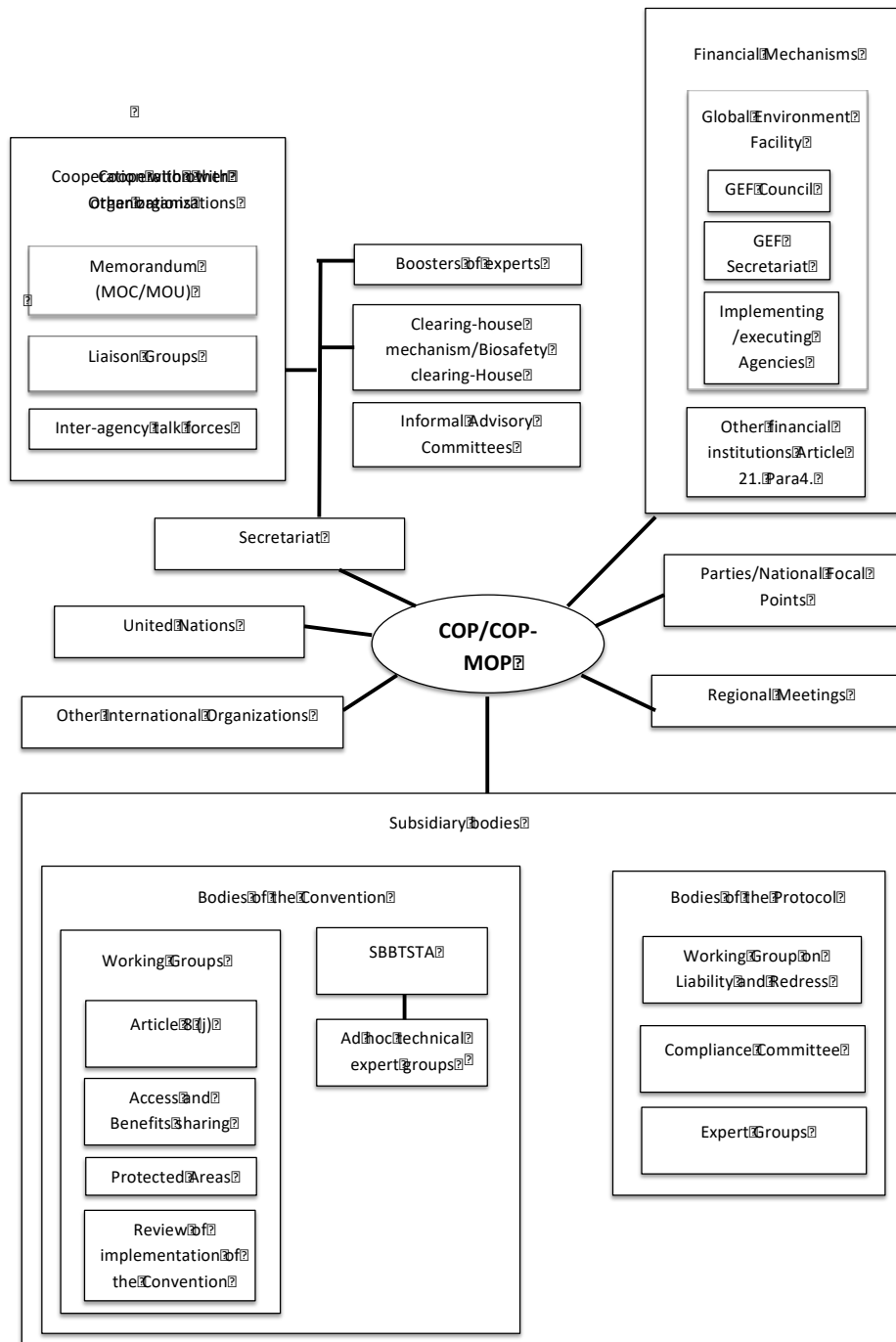


Figura 52. Estructura de los órganos de gobierno del CDB. Fuente: <http://www.cbd.int/cop>.

2.3.1. La Conferencia de las Partes (CoP)

Es el órgano rector integrado por los representantes de los Estados y organizaciones parte del CDB (Convention on Biological Diversity, s. f., CoP). Favorece los progresos del convenio mediante las decisiones que adopta en sus reuniones periódicas. Incluye también algunos observadores (ONG, Naciones Unidas o Estados que no son parte del convenio). Al principio, se reunía de forma anual, pero

desde 1998 se reúne cada dos años para examinar el progreso, fijar prioridades y adoptar planes de trabajo.

El importante papel atribuido a la CoP en los convenios relacionados con el medio ambiente y en el seguimiento de las acciones de las secretarías es una prueba de la preocupación que tienen los Estados por el control de los cambios normativos (Le Prestre, 2002c).

Las funciones de la CoP están recogidas en el artículo 23⁵⁵ del CDB y son las siguientes:

⁵⁵ Artículo 23. «Conferencia de las Partes 1. Queda establecida una Conferencia de las Partes. El Director Ejecutivo del Programa de las Naciones Unidas para el Medio Ambiente convocará la primera reunión de la Conferencia de las Partes a más tardar un año después de la entrada en vigor del presente Convenio. De allí en adelante, las reuniones ordinarias de la Conferencia de las Partes se celebrarán a los intervalos regulares que determine la Conferencia en su primera reunión. 2. Las reuniones extraordinarias de la Conferencia de las Partes se celebrarán cuando la Conferencia lo estime necesario o cuando cualquiera de las Partes lo solicite por escrito, siempre que, dentro de los seis meses siguientes de haber recibido de la secretaría comunicación de dicha solicitud, un tercio de las Partes, como mínimo, la apoye. 3. La Conferencia de las Partes acordará y adoptará por consenso su reglamento interno y los de cualesquiera órganos subsidiarios que establezca, así como el reglamento financiero que regirá la financiación de la Secretaría. En cada reunión ordinaria, la Conferencia de las Partes aprobará un presupuesto para el ejercicio financiero que transcurrirá hasta la reunión ordinaria siguiente. 4. La Conferencia de las Partes examinará la aplicación de este Convenio y, con ese fin:

- a) Establecerá la forma y los intervalos para transmitir la información que deberá presentarse de conformidad con el artículo 26,
- b) Examinará el asesoramiento científico, técnico y tecnológico sobre la diversidad biológica facilitado conforme al artículo 25;
- c) Examinará y adoptará, según proceda, protocolos de conformidad con el artículo 28;
- d) Examinará y adoptará, según proceda, las enmiendas al presente Convenio y a sus anexos, conforme a los artículos 29 y 30;
- e) Examinará las enmiendas a todos los protocolos, así como a todos los anexos de los mismos y, si así se decide, recomendará su adopción a las Partes en el protocolo pertinente;
- f) Examinará y adoptará anexos adicionales al presente Convenio, según proceda, de conformidad con el artículo 30;
- g) Establecerá los órganos subsidiarios, especialmente de asesoramiento científico y técnico, que se consideren necesarios para la aplicación del presente Convenio;
- h) Entrará en contacto, por medio de la Secretaría, con los órganos ejecutivos de los convenios que traten cuestiones reguladas por el presente Convenio, con miras a establecer formas adecuadas de cooperación con ellos; y
- i) Examinará y tomará todas las demás medidas necesarias para la consecución de los objetivos del presente Convenio a la luz de la experiencia adquirida durante su aplicación. 5. Las Naciones Unidas, sus organismos especializados y el Organismo Internacional de Energía Atómica, así como todo Estado que no sea Parte en el presente Convenio, podrán estar representados como observadores en las reuniones de la Conferencia de las Partes. Cualquier otro órgano u organismo nacional o internacional, ya sea gubernamental o no gubernamental, con competencia en las esferas relacionadas con la conservación y utilización sostenible de la diversidad biológica, que haya informado a la Secretaría de su deseo de estar representado, como

- a) Examinar la aplicación del Convenio.
- b) Dirigir, supervisar y decidir sobre el proceso de instrumentación, operación y futuro desarrollo del mismo con la asesoría de los órganos del CDB.

La CoP también puede crear grupos de trabajo sobre temas específicos abordados por el convenio. Un ejemplo es el grupo de trabajo sobre el artículo 8 (j).

2.3.2. La Secretaría Ejecutiva

Se creó para apoyar los objetivos del CDB. La CoP designó al PNUMA como la organización internacional competente para desempeñar las funciones de la Secretaría Ejecutiva.

Desde 1996 la sede de la Secretaría se encuentra situada en Montreal (Canadá). Su principal función es ayudar a los Gobiernos a aplicar el convenio y sus programas de trabajo, organizar reuniones, redactar borradores de documentos, coordinar la labor del convenio con la de otras organizaciones internacionales y recopilar y difundir información.

El secretario ejecutivo es el director de la Secretaría. También es responsable de las actividades de diseminación y de concienciación del público de conformidad con el artículo 13 del CDB.

El artículo 24⁵⁶ del CDB recoge las principales funciones de la Secretaría Ejecutiva, entre las que se encuentran:

observador, en una reunión de la Conferencia de las Partes, podrá ser admitido a participar salvo si un tercio, por lo menos, de las Partes presentes se oponen a ello. La admisión y participación de observadores estarán sujetas al reglamento aprobado por la Conferencia de las Partes». Disponible en <https://www.cbd.int/doc/legal/cbd-es.pdf>. (Consultado por última vez el 8 de mayo de 2015).

⁵⁶ Artículo 24: «Secretaría 1. Queda establecida una secretaría, con las siguientes funciones:

- a) Organizar las reuniones de la Conferencia de las Partes previstas en el artículo 23, y prestar los servicios necesarios;
- b) Desempeñar las funciones que se le asignen en los protocolos; 17
- c) Preparar informes acerca de las actividades que desarrolle en desempeño de sus funciones en virtud del presente Convenio, para presentarlos a la Conferencia de las Partes;
- d) Asegurar la coordinación necesaria con otros órganos internacionales pertinentes y, en particular, concertar los arreglos administrativos y contractuales que puedan ser necesarios para el desempeño eficaz de sus funciones; y
- e) Desempeñar las demás funciones que determine la Conferencia de las Partes.

2. En su primera reunión ordinaria, la Conferencia de las Partes designará la Secretaría escogiéndola entre las organizaciones internacionales competentes que se hayan mostrado dispuestas a desempeñar las funciones de Secretaría establecidas en el presente Convenio». Disponible en: <https://www.cbd.int/doc/legal/cbd-es.pdf>. (Consultado por última vez el 8 de mayo de 2015)

- Organizar las reuniones de la CoP y prestar los servicios necesarios.
- Desempeñar las funciones que se le asignen en los protocolos.
- Preparar informes acerca de las actividades que desarrolló en desempeño de sus funciones para presentarlos a la CoP.
- Asegurar la coordinación necesaria con otros órganos internacionales pertinentes, y en particular, concertar los arreglos administrativos y contractuales que puedan ser necesarios para el desempeño eficaz de sus funciones.
- Desempeñar otras funciones que sean determinadas por la CoP.

La Secretaría no tiene poder ni para decidir ni para aprobar normas de forma directa. Algunos observadores defienden que la Secretaría debería tener un rol más importante e independiente, habida cuenta de que la supervisión de que se está implementando de forma correcta lleva aparejada la capacidad de realizar una verificación independiente así como la capacidad de recopilar datos.

En general, los Estados tienden a fortalecer el rol tanto de la CoP como del Órgano Subsidiario de Asesoramiento Científico, Técnico y Tecnológico al que en adelante se hará referencia por sus siglas en inglés SBSTTA (Subsidiary Body on Scientific, Technological and Technical Advice).

2.3.3. *Órgano Subsidiario de Asesoramiento Científico, Técnico y Tecnológico*

Regulado en el artículo 25⁵⁷ del CDB, el SBSTTA nació como órgano científico de asesoramiento de la CoP, con vocación de replicar, de alguna manera, el órgano de

⁵⁷ Artículo 25. «Órgano subsidiario de asesoramiento científico, técnico y tecnológico 1. Queda establecido un órgano subsidiario de asesoramiento científico, técnico y tecnológico a fin de proporcionar a la Conferencia de las Partes y, cuando proceda, a sus otros órganos subsidiarios, asesoramiento oportuno sobre la aplicación del presente Convenio. Este órgano estará abierto a la participación de todas las Partes y será multidisciplinario. Estará integrado por representantes de los gobiernos con competencia en el campo de especialización pertinente. Presentará regularmente informes a la Conferencia de las Partes sobre todos los aspectos de su labor.

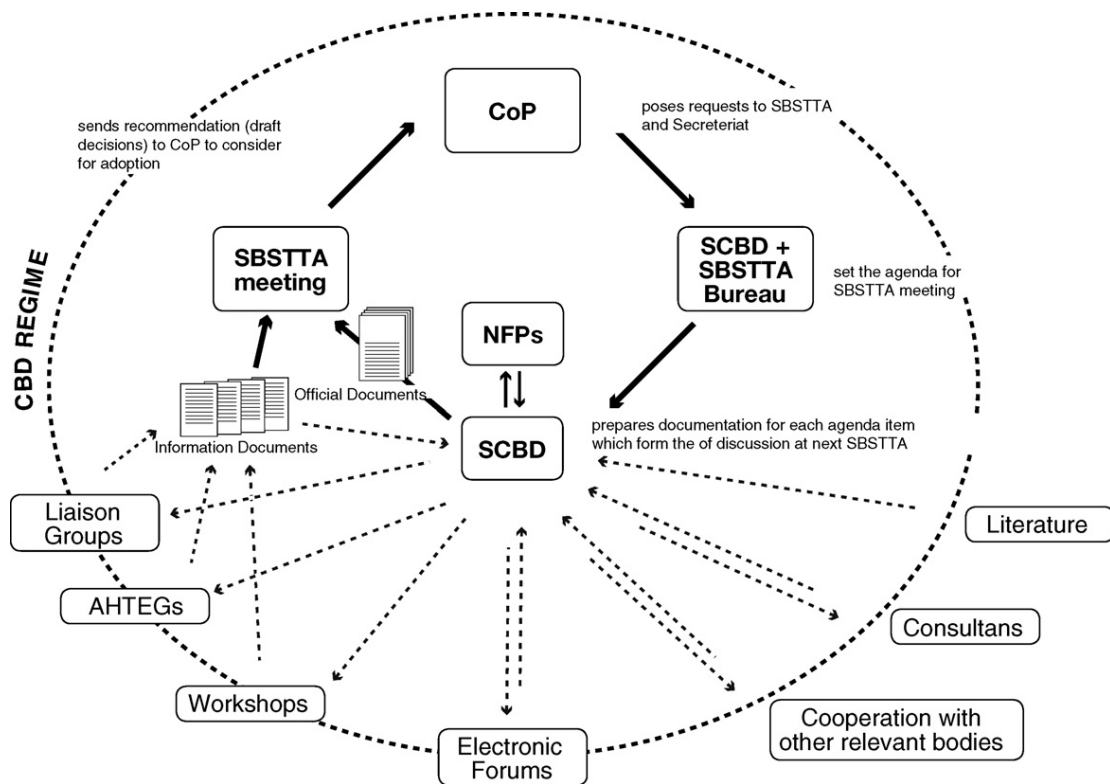
2. Bajo la autoridad de la Conferencia de las Partes, de conformidad con directrices establecidas por esta y a petición de la propia Conferencia, este órgano:

- a) Proporcionará evaluaciones científicas y técnicas del estado de la diversidad biológica;
- b) Preparará evaluaciones científicas y técnicas de los efectos de los tipos de medidas adoptadas de conformidad con las disposiciones del presente Convenio;
- c) Identificará las tecnologías y los conocimientos especializados que sean innovadores, eficientes y más avanzados relacionados con la conservación y la utilización sostenible de la diversidad biológica y prestará asesoramiento sobre las formas de promover el desarrollo o la transferencia de esas tecnologías;

expertos del IPCC. Es decir, con la expectativa de ser un órgano marcadamente científico, que permitiese la existencia de una interfaz científico-política. Recogiendo la definición (Van den Hove, 2007) estas interfaces se definen como procesos sociales que regulan las interacciones entre científicos y otros actores en el proceso de elaboración de políticas. Permiten realizar intercambios, coevolución y construcción conjunta del conocimiento con la finalidad de enriquecer la toma de decisiones.

La ciencia, prosigue Van den Hove (2007), es una actividad dentro de la sociedad cuya finalidad fundamental es generar conocimiento científico —conocimiento objetivo fundamentalmente en la forma de explicaciones del mundo y predicciones de su evolución—.

No hay duda de que para que exista un diálogo productivo entre ciencia y política debe ser posible generar objetivos que vinculen a las partes, desligados de los intereses «políticos» y liderados por la neutralidad del conocimiento.



- d) Prestará asesoramiento sobre los programas científicos y la cooperación internacional en materia de investigación y desarrollo en relación con la conservación y la utilización sostenible de la diversidad biológica; y 18
- e) Responderá a las preguntas de carácter científico, técnico, tecnológico y metodológico que le planteen la Conferencia de las Partes y sus órganos subsidiarios.

3. La Conferencia de las Partes podrá ampliar ulteriormente las funciones, el mandato, la organización y el funcionamiento de este órgano.» Disponible en: <https://www.cbd.int/doc/legal/cbd-es.pdf> (Consultado por última vez el 8 de abril de 2017).

Figura 53. Procesos intersesionesales del SBSSTA. Fuente: Koezt *et al.*, 2008.

Sin embargo, el SBSSTA es un órgano político y existen buenas razones para que continúe siéndolo (Koezt *et al.*, 2008), aunque esto no es óbice para que tenga un mejor funcionamiento como interfaz científico-política.

La cuestión no es tanto cómo hacer que el SBSSTA sea un órgano más científico, sino de qué manera puede proporcionar una opinión más técnica y creíble, eso sí, desde el punto de vista científico. Aunque, en principio, lo político y lo científico parecen estar en controversia, no es así.

Si la politización de la ciencia y la «cientificación» de la política son hasta cierto punto aspectos a tener en cuenta en lo que respecta a la intersección de ciencia y política, algunas intersecciones profundas, ocultas, pueden conducir a la confusión y a la inestabilidad (Guston, 2001).

En efecto, de conformidad con la Secretaría: «el SBSSTA tiene un importante papel que desempeñar haciendo de puente entre la comunidad científica y los que redactan las normas» (CBD, 1999),⁵⁸ en particular, mediante la construcción de vínculos de cooperación con la comunidad científica (Le Prestre, 2002c).

Desde el mismo momento de su creación, existió una tensión en relación con el papel que debía desempeñar. Esta tensión era un reflejo de las distintas ideas en relación con las prioridades del convenio. En efecto, mientras algunos Estados eran partidarios de potenciar el aspecto de conservación del CDB (por ejemplo, Francia), otros querían un SBSSTA profundamente científico que apoyase todas las actividades y otros defendían la necesidad de evolucionar y de aproximarse a la estructura del IPCC aplicándola a biodiversidad.

También existía una desconfianza en el rol que iba a desempeñar con base en distintos factores (Vadrot, 2014), entre los cuales los más relevantes eran los temas de soberanía nacional, el acceso a la biodiversidad y el acceso a sus potenciales resultados. Además, los países en desarrollo eran contrarios a la creación del SBSSTA, por estar en clara desventaja al no poder aportar tanto conocimiento científico como los países desarrollados (Karlsson, 2007). De alguna manera, las concesiones «excesivas» realizadas en el ámbito del IPCC llevaron a los firmantes del CDB a limitar los compromisos.

⁵⁸ Disponible en <https://www.cbd.int>. (Consultado por última vez el 23 de enero de 2017).

2.4. ¿El fracaso del SBSTTA?

No cabe duda de que el SBSTTA no está siendo lo efectivo que debiera. En efecto, aunque concebido como un mecanismo de asesoramiento con un marcado carácter científico, se ha ido separando del concepto inicial. Incluso el propio Koetz, sin cuestionar la propia existencia del SBSSTA, reconoce que precisa ser más creíble (Koezt *et al.*, 2008).

Como no puede ser de otra manera, la ciencia está en la base de todas las convenciones de biodiversidad. Por lo tanto, resulta preciso realizar una evaluación científica. Sin embargo, el SBSSTA no ha servido de catalizador de nuevos programas de investigación ni de espacio compartido entre científicos y políticos (Le Prestre y Compagnon, 2016).

Es más, muchos autores ya no hablan de SBSTTA, sino de mini CoP (Vadrot, 2014), es decir de un órgano con marcados rasgos políticos que negocia borradores de decisiones preparando las subsiguientes reuniones de la CoP (Koezt *et al.*, 2008).

Uno de sus principios fundamentales era mejorar de forma continuada la calidad de su asesoramiento, mejorando los *inputs* técnicos y tecnológicos usados en el debate y trabajando en las reuniones del SBSTTA (Koezt *et al.*, 2008). Con carácter general, se considera que la interfaz entre la ciencia y la política debe reforzarse fundamentalmente en el ámbito internacional con la finalidad de lograr una mejor gobernanza de la biodiversidad (Loreau y Oteng, 2006).

Sin embargo y a pesar de los esfuerzos realizados para dar un carácter más científico y menos político al SBSTTA, sigue siendo una pre-CoP, en la que las partes negocian las ramificaciones normativas de la mejor ciencia disponible (Görg, 2010). Algunos países rechazaron las evaluaciones científicas independientes como pueden ser la Evaluación de la Biodiversidad Mundial debido a una falta de legitimidad política, por lo que tuvo escaso impacto en los procesos de gobernanza (Heywood y Watson, 1995).

Todas estas reflexiones han contribuido al debate existente sobre si el SBSTTA tiene que ser modificado para proporcionar estrictamente recomendaciones científicas, o bien mantener el rol político que tiene en este momento y seguir negociando las decisiones sustantivas de la CoP. Se ha llegado incluso a plantear si el SBSTTA tiene capacidad para elaborar sólidas recomendaciones científicas y técnicas (Koezt *et al.*, 2008).

Mientras tanto, sigue atrapado y tratando de convertirse en el cuerpo de carácter científico con cuya vocación nació y dejar a un lado la naturaleza más política que ha desarrollado, en la realidad, un carácter político que oficialmente no puede tener. Un

indicativo del «carácter no científico» de este órgano es que solo el 7 % de los integrantes proceden directamente del ámbito académico o de investigación.

Dos son los reproches fundamentales que se le hacen: por una parte, las insuficientes aportaciones científicas en las reuniones y, por otra, la calidad de los debates. El poder y el prestigio de la ciencia descansan en la capacidad de los científicos de separar los intereses científicos de los no científicos (Moore, 1996). Dado que el SBSTTA no es políticamente neutro, resulta difícil diferenciarlo de la CoP. Además, esta es reticente a darle autonomía, con lo que demuestra su afán por no permitir que un órgano científico adopte decisiones vinculantes o presente a la CoP *faits accomplis* en términos de fijación de agenda (Le Prestre, 2002).

Así, habida cuenta de la naturaleza del SBSTTA, sus antecedentes y el hecho de que son las partes las que deciden a quién enviar, no es extraño que las personas que toman parte en las reuniones rara vez sean elegidas teniendo en cuenta sus referencias técnicas y científicas en relación con el tema objeto de discusión. Es un órgano político donde van embajadores y otros diplomáticos, en lugar de un contexto científico donde tomar decisiones. Los Estados no quieren perder soberanía sobre estos temas.

De alguna manera, tal vez, la falta de ciencia o de órganos científicos en el CDB es uno de los motivos más repetidos para explicar que no se hayan conseguido los resultados esperados en términos de frenar la pérdida de biodiversidad. El proceso preveía una reunión del SBSTTA, supuestamente científica, en los tres meses anteriores a la celebración de la reunión de la CoP. Dicha reunión serviría de base para la discusión de la CoP, con una neutralidad política y un enfoque eminentemente científico. No ha funcionado y los objetivos del CDB están lejos de haberse conseguido.

Por lo tanto, sigue siendo necesaria una plataforma global comparable con el IPCC respecto al cambio climático para dar una respuesta global coordinada en materia de pérdida de biodiversidad. Un informe elaborado en 2010 (Secretaría del Convenio sobre la Diversidad Biológica, 2010) puso de manifiesto que la revisión de ese año del plan estratégico para el CDB constituía una oportunidad para definir esa visión y fijar metas con plazos específicos que fomentasen la adopción de las medidas necesarias para alcanzarla. En efecto, el análisis de 2010 demostró que no se habían alcanzado las metas establecidas. A pesar de ser un convenio casi universal, no fue capaz de impulsar la adopción de las medidas necesarias. Eso hace necesario cambiar la dirección.

Así, la gestión de la gobernanza internacional requiere de la creación de mejores interfaces científico-políticas (Koetz y Bridgewater, 2012). Esta es una de las razones para que nazca la Plataforma Intergubernamental sobre Biodiversidad y Servicios de

los Ecosistemas (IPBES, por las siglas en inglés de Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services). En efecto, una de las claves para explicar que no haya podido frenarse la pérdida de biodiversidad ha sido la inexistencia de interfaces adecuadas (Koetz y Bridgewater, 2012).

Por una multiplicidad de razones, el SBSSTA no ha sido capaz de afrontar los retos científicos que se le plantearon. La justificación de la politización de la que hablan los científicos no es la única razón. Los científicos tampoco han sabido hacer llegar al órgano las necesidades reales, conseguir el compromiso. Por lo tanto, el entramado organizaciones del CDB no ha permitido que cumpliera los objetivos para los que nació.

Otra de las reflexiones que se ha planteado es que el CDB nació para ser un instrumento imperativo de derecho internacional, con voluntad de tener efectos legales. Su estructura o marco que le sirvió para arrancar ha hecho que se quedara en un instrumento dispositivo (Harrop y Pritchard, 2011).

2.5. El CDB y los objetivos de frenar la pérdida de diversidad biológica

Sin embargo, es un hecho que el CDB no ha acabado de conseguir su objetivo de frenar la pérdida de biodiversidad. Existen una multiplicidad de razones que explican que no se hayan alcanzado los objetivos. Tal vez, el CDB no haya contado con los mecanismos necesarios para lograr el compromiso suficiente de las partes. Tal vez, la implementación no haya sido la adecuada o la politización del SBSSTA no haya contribuido a la consecución de los objetivos.

En todo caso, es evidente que el objetivo de frenar la pérdida de biodiversidad no se ha conseguido. Una de las razones fundamentales es la falta de consenso a la hora de medir la pérdida de biodiversidad. Sin lugar a dudas, merece una consideración especial el consenso que muchos autores consideran clave en la gobernanza de la biodiversidad, pero la superposición de intereses y de *stakeholders* diversos complica mucho lograrlo (Morin *et al.*, 2016).

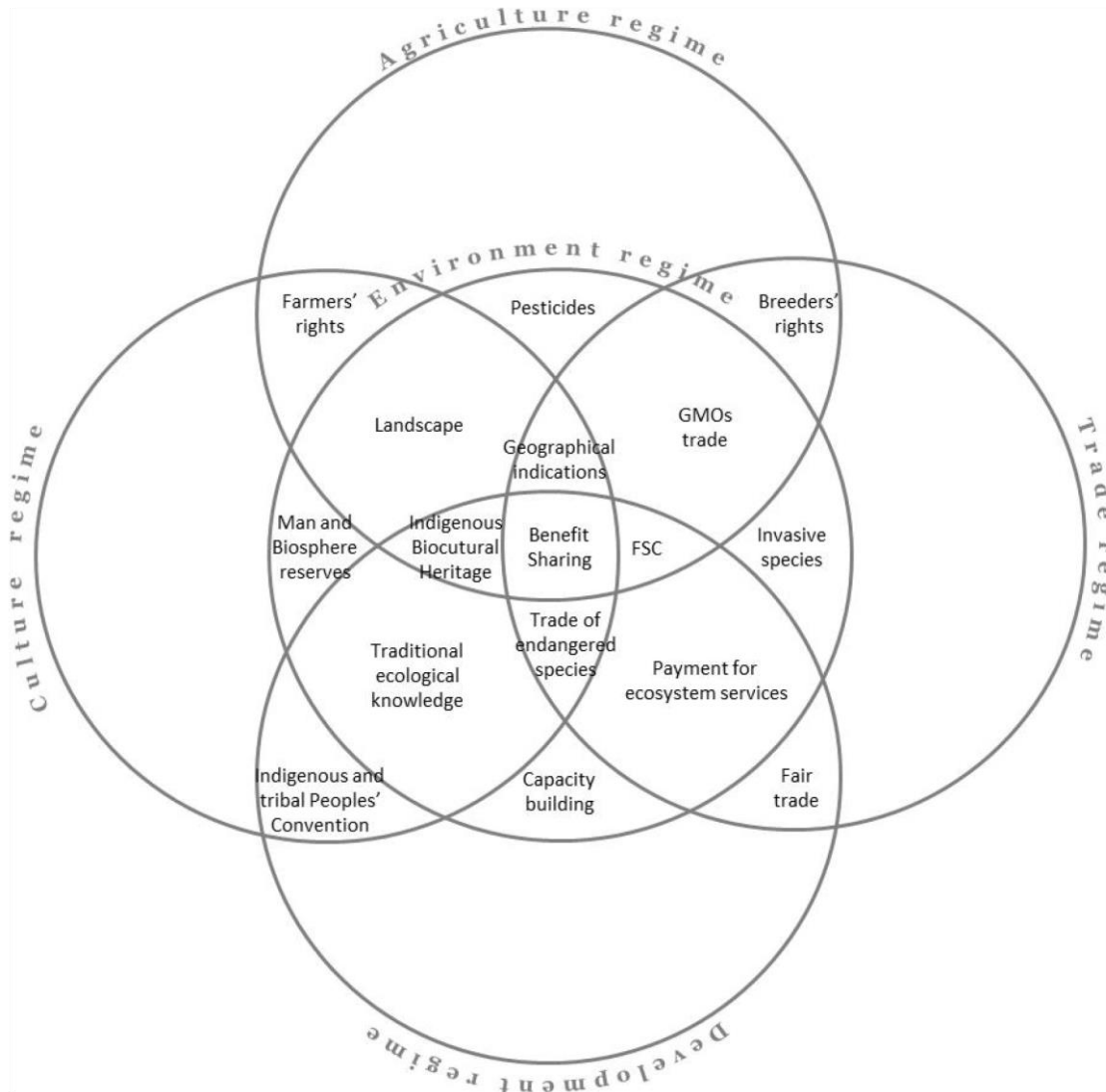


Figura 54. La complejidad de la biodiversidad. Fuente: Morin, Louafi, Orsini y Oubenal, 2018.

Otra razón que ha dificultado conseguir los objetivos es la falta de desarrollo o inexistencia de los indicadores necesarios de biodiversidad por parte de los que realizan las políticas. Tampoco se ha invertido en ellos lo suficiente como para poderlos desarrollar (Walpole *et al.*, 2009).

Existe la necesidad de que el CDB promueva la generación de indicadores de biodiversidad y de servicios de los ecosistemas rigurosos, repetibles, ampliamente aceptados y que puedan comprenderse fácilmente (Balmford y Bennun, 2005), pues no existen unidades de medida que permitan calcular la pérdida de biodiversidad. La necesidad de replicar el modelo del IPCC en el marco de la biodiversidad sigue presente.

En 2010, los miembros del CDB adoptan el Plan Estratégico para la Biodiversidad (2011-2020), con la misión de parar la pérdida de biodiversidad y potenciar los beneficios que ello supone para las personas. Con esa finalidad, se hacen los

llamados Objetivos Aichi de Biodiversidad (Aichi Targets) organizados en cinco objetivos estratégicos.

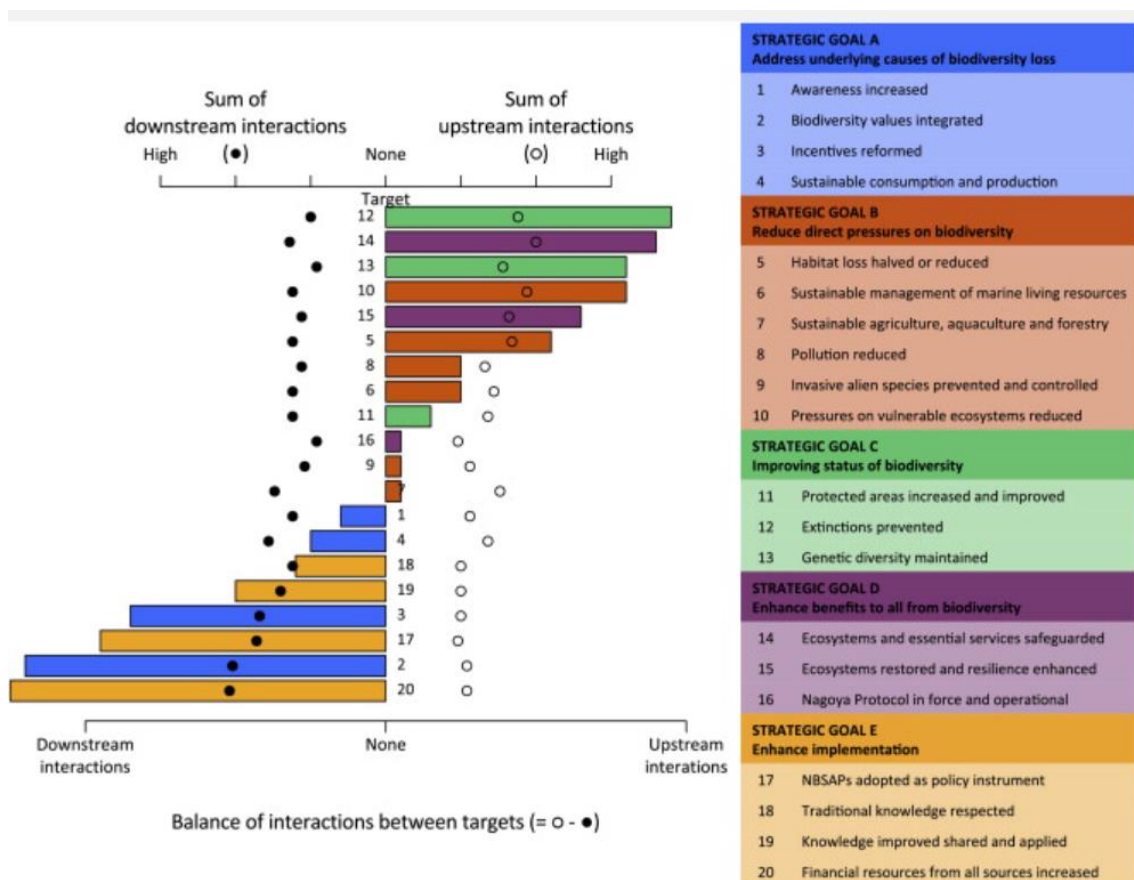


Figura 55. Interacciones entre los Aichi Targets. A la izquierda, pueden verse las interacciones netas de los distintos objetivos (en barras) medidos respecto a la diferencia entre la suma de las interacciones descendentes que es igual a la suma de los impactos que afectan a los demás objetivos, y la suma de las interacciones ascendentes, que es igual a la suma de los impactos recibidos de los demás objetivos. Las acciones dirigidas a los impactos sobre los objetivos.

En el contexto del CDB y con la finalidad de establecer indicadores para medir la pérdida de diversidad, se crean los Aichi Targets. Se definen como un conjunto de objetivos globales⁵⁹ estructurados dentro del Plan Estratégico de Biodiversidad 2011-2020. Están agrupados dentro de cinco objetivos estratégicos:

1. Abordar las causas subyacentes de la pérdida de biodiversidad mediante la integración de la biodiversidad en el Gobierno y la sociedad.
2. Reducir las presiones directas sobre la biodiversidad y promover el uso sostenible.

⁵⁹ Recuperado de <http://biodiversitya-z.org/content/aichi-biodiversity-targets>. (Consultado por última vez el 14 de marzo de 2018).

3. Mejorar el estado de la biodiversidad salvaguardando los ecosistemas, las especies y la diversidad genética.
4. Mejorar los beneficios para todos de la biodiversidad y los servicios de los ecosistemas.
5. Mejorar la implementación a través de la planificación participativa, la gestión del conocimiento y la creación de capacidades.

Cada uno de los objetivos recogido en el Plan Estratégico aborda un reto diferente en relación con la preocupación de parar la pérdida de biodiversidad. El primero de los objetivos aborda la necesidad de cambios socio-económicos e institucionales. El segundo de ellos, se centra en reducir las presiones directas que existen en materia de biodiversidad y de ecosistemas. El tercero abarca enérgicos esfuerzos en mejorar el status de la biodiversidad. El cuarto persigue garantizar el flujo de beneficios de la biodiversidad y de los ecosistemas a las personas, fundamentalmente en comunidades cuya existencia tiene una fuerte vinculación con servicios de los ecosistemas locales. Para terminar, el quinto persigue desarrollar las condiciones necesarias para la implementación del Plan Estratégico así como el desarrollo de la base del conocimiento (Marques, A. *et al.* 2014).

Aunque la iniciativa de los Aichi Targets fue un gran paso, la realidad es que estamos muy lejos de haberlo conseguido. Una de las causas principales, sin duda, es la multiplicidad de factores que influyen en materia de biodiversidad y la dificultad que existe a la hora de definir con la suficiente claridad indicadores que permitan delimitar, de verdad, la pérdida de biodiversidad. Muestra de ello es en la indiscutible ambigüedad con la que se han formulado los objetivos que hace tan difícil su interpretación (Butchart, S.H.M., *et al.* 2016). A esta dificultad, se añade la inexistencia de elementos que permitan cuantificarlos. No parece que tenga sentido establecer unas referencias que no tengan aparejados indicadores claros para medirlos. Otros factores adicionales como la complejidad o la redundancia dificultan la tarea. Por eso, concluyen que los futuros objetivos de biodiversidad deben ser específicos, simples, sucintos, cuantificables, sin ambigüedad, limitados en número y fijados a través de un proceso que suponga una mayor colaboración entre los científicos y los que redactan las políticas (Butchart, S.H.M., *et al.* 2016).

Posiblemente una de las complicaciones mayores en materia de biodiversidad es llegar a definir que es suficiente y eficiente a la hora de elaborar estrategias de conservación (Di Marco *et al.* 2016).

Aunque hoy se tiene más conocimiento sobre el estado de la biodiversidad (Rands *et al.*, 2010), sobre los procesos biológicos y sociales que le afectan y sobre las presiones y *drivers* subyacentes que promueven su continuo declive, siguen existiendo vacíos de conocimiento. De hecho, el conocimiento que hoy tenemos en

relación con la gestión de ecosistemas y su restauración es inadecuado para la consecución de los retos.

La evaluación de los ecosistemas exige el desarrollo y la aplicación de indicadores adecuados. Tienen que ser (i) confiables y capaces de simplificar relaciones complejas, (ii) cuantificables y transparentes para facilitar la comunicación y (iii) adecuados para el propósito con el que se implementen (Feld *et al.*, 2010).

Sin lugar a dudas, la toma de decisiones adecuadas requiere la adopción de indicadores de biodiversidad globales (Nicholson *et al.*, 2012).

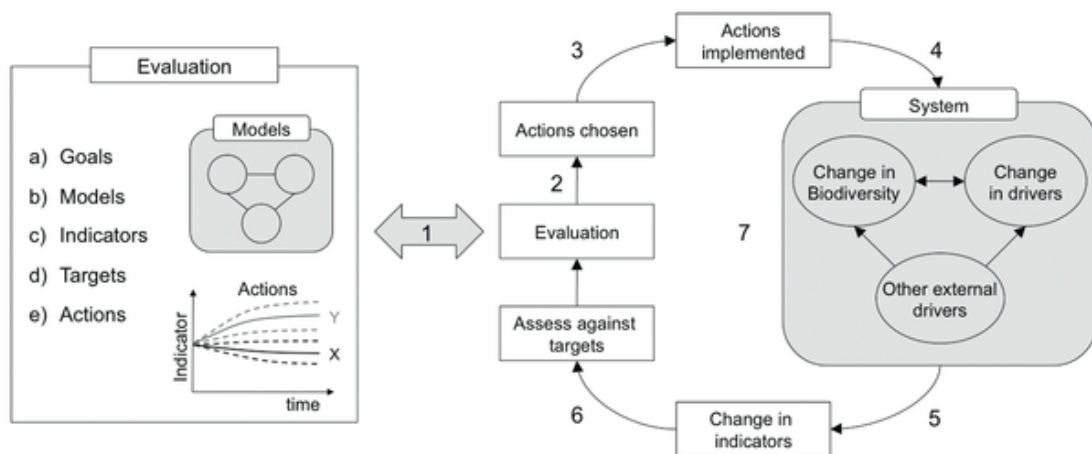


Figura 56. Políticas de indicadores de ciclos. Marco para usar indicadores para informar la política: la evaluación del problema implica definir objetivos generales de política, desarrollar modelos del sistema, seleccionar indicadores que reflejen los cambios de interés, definir objetivos específicos que puedan medirse con los indicadores y definir un conjunto de acciones o políticas para lograr los objetivos, evaluados con los modelos. Tras la evaluación, se eligen e implementan acciones, con el consiguiente cambio en la biodiversidad y en los factores de pérdida, que también pueden verse afectados por otros factores externos no relacionados con las acciones; el cambio se supervisa con los indicadores y se evalúa en relación con los objetivos, y ello lleva a una reevaluación. Las principales fuentes de incertidumbre y posibles fallas a lo largo del ciclo se numeran y discuten en detalle en la tabla S1: 1) supuestos en el proceso de evaluación; 2) vínculo entre evaluación y selección de acciones; 3) vínculo entre la selección y la implementación de acciones; 4) impacto de la acción que difiere del impacto anticipado; 5) vínculo entre cambio de biodiversidad y cambio de indicador; 6) vínculo entre el cambio de indicador y la evaluación del objetivo y 7) desajustes en escalas temporales y espaciales a lo largo del ciclo. Fuente: Nicholson *et al.* 2012.

Por lo tanto, la creación de los indicadores necesarios queda pendiente de resolver. Este es el punto de arranque de la Plataforma Intergubernamental sobre Biodiversidad y Servicios de los Ecosistemas (IPBES, por las siglas en inglés de

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services).⁶⁰

3. LA PLATAFORMA INTERGUBERNAMENTAL SOBRE BIODIVERSIDAD Y SERVICIOS DE LOS ECOSISTEMAS

La biodiversidad de los ecosistemas terrestres, marinos, costeros y de aguas continentales sirve de base para los ecosistemas y los servicios que sostienen el bienestar humano. No obstante, la biodiversidad y los servicios de los ecosistemas están disminuyendo a un ritmo sin precedentes y, con la finalidad de hacer frente a este reto, es necesario adoptar y desarrollar políticas adecuadas en el ámbito local, nacional e internacional.

Para lograrlo, los que adoptan las decisiones necesitan información científicamente creíble e independiente que tenga en cuenta la complejidad de las relaciones entre la biodiversidad, los servicios de los ecosistemas y las personas. Necesitan también métodos efectivos para interpretar esta información científica con el objeto de poder tomar decisiones informadas. La comunidad científica tiene que comprender asimismo las necesidades de los que toman las decisiones para proporcionarles la información relevante. Fundamentalmente, el diálogo entre la comunidad científica, los gobiernos y otros interlocutores en materia de biodiversidad y de servicios de ecosistemas debe fortalecerse y legitimarse (PNUMA, 2009).

El Grupo Intergubernamental de Expertos sobre el Cambio Climático, conocido como IPCC (*Intergovernmental Panel on Climate Change*), es una organización internacional, constituida a petición de los Gobiernos miembros. Fue establecido por primera vez en 1988 por dos organizaciones de Naciones Unidas, la Organización Meteorológica Mundial (OMM) y el Programa de las Naciones Unidas para el Medio Ambiente (PNUMA), y posteriormente ratificada por la Asamblea General de las Naciones Unidas. El funcionamiento del IPCC, su eficacia y la falta de resultados en materia de biodiversidad han hecho que los expertos en biodiversidad miren y traten de reproducir este modelo con la creación del IPBES. El IPCC es un organismo internacional que coordina la ciencia ajeno a la estructura del CDB.

La resolución por la que se establece el IPBES enfatiza, asimismo, la necesidad de realizar asesoramientos a todos los niveles: global, regional y en niveles inferiores al regional teniendo en cuenta los *inputs* de los *stakeholders* a distintos niveles, incluidas

⁶⁰ Para consultar más información sobre la plataforma, véase: www.ipbes.net/about-ipbes.html. (Consultado por última vez el 12 de julio de 2016).

las organizaciones científicas, organizaciones no gubernamentales, comunidades locales y negocios (PNUMA, 2012).

Más allá de todo esto, se supone que IPBES tiene que configurarse como un vínculo entre la ciencia y la toma de decisiones informadas, incorporando todas las disciplinas relevantes y teniendo los distintos sistemas de conocimiento, incluidos los sistemas indígenas, para construir una interfaz científico-política (PNUMA, 2012).

Por eso, proporciona un mecanismo que, al mismo tiempo, reconocen las comunidades científica y política para sintetizar, revisar, asesorar y evaluar de forma crítica información relevante y el conocimiento en el ámbito mundial por parte de Gobiernos, académicos, organizaciones científicas, organizaciones no gubernamentales y comunidades indígenas.

Es un grupo de expertos con credibilidad para realizar asesoramientos con esta información y conocimiento de una forma transparente. Es único en su objetivo de reforzar la capacidad para el uso efectivo de la ciencia en la toma de decisiones a todos los niveles. Tiene también como finalidad satisfacer las necesidades los acuerdos medioambientales multilaterales relacionados con la biodiversidad y los servicios de los ecosistemas y trabajar en los procesos que existen para asegurarse de que se crean sinergias y que los trabajos de cada una de ellas se complementan. Además, el IPBES es una red gigante donde convergen Gobiernos, organizaciones, científicos, etc.; las bases están puestas.

3.1. Los orígenes de IPBES

Se habla del fracaso del CDB, por no haberse conseguido los objetivos de parar la pérdida de biodiversidad perseguidos (Van Den Hove y Chabason, 2009). Importantes voces dentro de la comunidad científica ponen de manifiesto que, con la finalidad de afrontar la cuestión de la biodiversidad, es de crítica importancia hacer una desconexión entre lo científico y lo político y crear una organización permanente, gestionada de forma global similar al IPCC (Loreau y Oteng, 2006; Diversitas, 2005).

Al hilo de ello, continúan haciéndose esfuerzos con la finalidad de mejorar los procesos de evaluación de la biodiversidad, siempre tomando como punto de referencia el IPCC, todo un éxito en el ámbito del cambio climático.

Así, a finales de los años noventa, surge la Evaluación de los Ecosistemas del Milenio (MEA, por las siglas en inglés de Millenium Ecosystem Assessment)

(Larigauderie y Mooney, 2010).⁶¹ Tiene su origen en movimientos de científicos y actores políticos e introduce un nuevo marco para analizar los sistemas sociales y ecológicos que tienen una amplia influencia tanto en las comunidades científicas como en las de gobierno (Carpenter *et al.*, 2009).

A mediados de los noventa muchas personas involucradas en convenciones internacionales como el CDB y la Convención de Lucha contra la Desertificación (CLD) (Naciones Unidas, 2007) comprenden que las extensivas necesidades de evaluación científica en las convenciones no están siendo satisfechas por los mecanismos establecidos. En cambio, en ámbitos distintos al de la biodiversidad, existen procesos de evaluación efectivos, como IPCC, y tratados como la Convención Marco sobre el Cambio Climático (Naciones Unidas, 1992) y el Protocolo de Montreal sobre Sustancias que Agotan la Capa de Ozono (Programa de las Naciones Unidas para el Medio Ambiente, 2000).

Los científicos también identifican la necesidad de una evaluación internacional de los ecosistemas. A pesar de existir grandes avances en las ciencias ecológicas, en la economía de los recursos naturales y en otros ámbitos durante los años ochenta y noventa, los nuevos avances aparecen pobremente reflejados en las discusiones de políticas sobre los ecosistemas. Reconociendo estas deficiencias, un panel de cuarenta científicos líderes prepara un borrador de evaluación internacional — *Proteger nuestro Planeta, Asegurar nuestro Futuro: Vínculos entre las Cuestiones Ambientales Globales y las Necesidades Humanas*—.

La Evaluación de los Ecosistemas del Milenio aporta una innovación a las evaluaciones de diversidad biológica (Larigauderie y Mooney, 2010). Incluye las ciencias sociales y naturales y se centra no solo en el estatus de los elementos de la biodiversidad, sino también en sus productos y funciones, fundamentalmente en servicios a los ecosistemas o beneficios que recibe la sociedad resultante del funcionamiento de los ecosistemas. Esta forma de hacer las cosas resulta ser un éxito y demuestra que el 60 % de los servicios que proporcionan los ecosistemas se han degradado durante los últimos cincuenta años.

Aunque se plantea como un proceso puntual, queda claro todo lo que queda por hacer. Debe hacerse un seguimiento del proceso para consolidar y actualizar tanto el concepto como los hallazgos. Es un hito fundamental en la génesis de IPBES (Hrabanski y Pesche, 2015).

Veinte años después del desarrollo del CDB, la creación de IPBES supone la culminación de un largo proceso destinado a poner la biodiversidad en la agenda

⁶¹ Esta historia se puede consultar en <http://www.millenniumassessment.org/es/History.html>. (Consultado por última vez el 14 de marzo de 2018).

internacional y a conseguir la institucionalización de un permanente proceso de evaluación del conocimiento que representa un logro que muchas veces se compara con el éxito del IPCC (Hrabanski y Pesche, 2015).

El objetivo de IPBES es reforzar la interfaz científico-político para la biodiversidad y los servicios de los ecosistemas con la finalidad de conservar la biodiversidad y de que tenga un uso sostenible, de lograr el bienestar humano a largo plazo y un desarrollo sostenible (Granjou *et al.*, 2013).

La idea de crear una organización global encargada de la biodiversidad y de los servicios de los ecosistemas nació en el mundo científico. Sin embargo, muy rápidamente los científicos se volvieron al PNUMA (Granjou *et al.*, 2013). El IPBES fue creado para cubrir las lagunas previamente identificadas por los expertos en políticas y los propios científicos (Vadrot, 2014).

En la creación de IPBES Granjou *et al.* (2013) hacen referencia a tres fases:

1. La fase de consulta

El precursor de IPBES fue el International Mechanism of Scientific Expertise on Biodiversity (IMoSEB, por sus siglas en inglés). De esta reunión salió la Declaración de París (Foro de Alto Nivel, 2005). La consulta se inició mediante una serie de casos de estudio utilizados para documentar las necesidades y para proponer alternativas, discutidas después en consultas regionales, con cientos de *stakeholders*.

La consulta IMoSEB terminó con un *Steering Committee* en el que se requirió al director ejecutivo del PNUMA para que convocara una reunión intergubernamental invitando a Gobiernos relevantes y a organizaciones no gubernamentales con la idea de crear una interfaz científico-política internacional eficiente; intergubernamental, pero también con *stakeholders* no gubernamentales, construyendo sobre la Evaluación de los Ecosistemas del Milenio. De esta manera, la informal IMoSEB se convirtió en un marco formal del PNUMA, donde empezó a fraguarse la negociación del IPBES.

2. La fase diplomática

En 2007, los líderes de IMoSEB decidieron celebrar una reunión intergubernamental e invitar al PNUMA. En la reunión celebrada en 2008, el proyecto IMoSEB evolucionó al proyecto del IPBES que debía alinearse con el seguimiento estratégico del MEA (Millennium Ecosystem Assessment, 2005).

Se celebraron tres reuniones intergubernamentales con una multiplicidad de interlocutores (Malasia 2008, Kenia 2009, República de Corea 2010) y con la finalidad de fortalecer una interfaz científica y normativa en biodiversidad y servicios de los ecosistemas. En las dos primeras reuniones se identificaron las

lagunas y la necesidad de reforzar la interfaz científica y normativa. En la reunión de junio 2010, en Busán (República de Corea), los Gobiernos decidieron crear el IPBES. El objetivo de su programa de trabajo debía elaborarse y la mayor parte de sus principios acordarse en la reunión de Busán, y ser parte de los resultados de esta reunión.

El documento final de Busán (PNUMA, 2010) fue presentado en la 10.^a Conferencia de las Partes del CDB en Nagoya en octubre de 2010 y fue analizado en la sesión número 65 de la Asamblea de Naciones Unidas. La Asamblea aprobó una resolución en la que solicitaba al PNUMA que convocara una reunión plenaria para el pleno funcionamiento del IPBES en cuanto tuviera oportunidad. El PNUMA aplicó esta resolución en una decisión en la sesión número 26 del Consejo de Administración, celebrada en febrero de 2011.

3. La fase de implementación

Las Naciones Unidas anunciaron la decisión de crear el IPBES oficialmente en la asamblea plenaria del final del «Año de la Biodiversidad», es decir, en diciembre de 2010. Planificaron una reunión para que el IPBES estuviera plenamente operativo a principios de octubre de 2011, con la idea de estructurar y organizar la forma de gobierno, y una segunda reunión para 2012, con la finalidad de elaborar el borrador del plan de trabajo.

La reunión plenaria se celebró en dos sesiones, la primera del 3 al 7 de octubre en Nairobi y la segunda auspiciada por el PNUMA, en colaboración con la UNESCO, la FAO y el PNUD, en la ciudad de Panamá del 16 al 21 de abril de 2012. Allí, muchas de las modalidades y acuerdos institucionales para la plataforma fueron completados y noventa y cuatro Gobiernos adoptaron una resolución que establecía la plataforma como un organismo intergubernamental independiente.

La primera reunión del Pleno del IPBES se celebró en Bonn (Alemania), los días 21 a 26 de enero de 2013, auspiciada por el Gobierno de Alemania. El documento final de esta sesión puede encontrarse en IPBES/1/12 (Naciones Unidas, 2013). Recoge decisiones en relación con los siguientes pasos para el desarrollo de un programa de trabajo inicial: el estado de las contribuciones y el presupuesto inicial de la plataforma para 2012, los arreglos administrativos e institucionales de IPBES y el procedimiento para recibir y priorizar las peticiones incluidas en la plataforma. Además, el informe incluye las reglas actualizadas del procedimiento para el pleno de la plataforma.

3.2. El nacimiento de IPBES

Formalmente y como se ha visto en el apartado anterior, IPBES nació en 2012 como un órgano intergubernamental abierto a todos los miembros de Naciones Unidas con la finalidad de «fortalecer la interfaz entre lo científico y lo político en lo que respecta a la biodiversidad y a los servicios de los ecosistemas para la conservación y uso sostenible de la biodiversidad, para el bienestar de las personas a largo plazo y para el desarrollo sostenible» (IPBES, s. f.).

Se creó con la misión de controlar la aplicación del CDB hasta 2020, priorizando el interés y las capacidades del sector privado sobre la utilización sostenible de la biodiversidad y de los servicios de los ecosistemas como fuente de futuras operaciones comerciales y como condición de nuevas posibilidades comerciales (Vergara, 2011).

En abril de 2012, el organismo quedó formalmente instituido en Panamá como una nueva plataforma para fortalecer la interfaz científico-normativa sobre la diversidad biológica terrestre, marina y de aguas interiores y sus interacciones, así como los servicios de los ecosistemas.

Aunque existen una multiplicidad de organizaciones e iniciativas que contribuyen a mejorar el diálogo entre los formuladores de políticas y la comunidad científica en este campo, el IPBES se estableció como una nueva plataforma reconocida tanto por la comunidad política como por la comunidad científica para abordar las brechas existentes. Se espera que se convierta en el punto focal clave de todas las agencias y las organizaciones involucradas en la conservación y el uso sostenible de la diversidad biológica.

La naturaleza de esfuerzos de evaluación como los de IPBES y lo que esta va a aportar resulta complejo de analizar. Para tener éxito como interfaz científico-política en relación con la biodiversidad y los servicios de los ecosistemas a nivel internacional deben reunirse cuatro elementos, cada uno complementario del otro: la investigación, la observación, la evaluación y la gobernanza (Larigauderie y Mooney, 2010). Sin la correcta implantación de estos cuatro elementos, la interfaz no puede funcionar de forma adecuada.

El trabajo del IPBES fue innovador en dos aspectos (Díaz *et al.*, 2015). En primer lugar, fue construido de una forma transparente, inclusiva y participativa, mediante numerosas reuniones y grupos de trabajo multidisciplinares y estuvo abierto a la revisión de un gran número de países y de agentes durante más de dos años. En segundo lugar, de forma expresa reunió varias disciplinas científicas (ciencias sociales, naturales, ingenierías, etc.) así como diversos agentes y sus distintos sistemas de conocimiento.

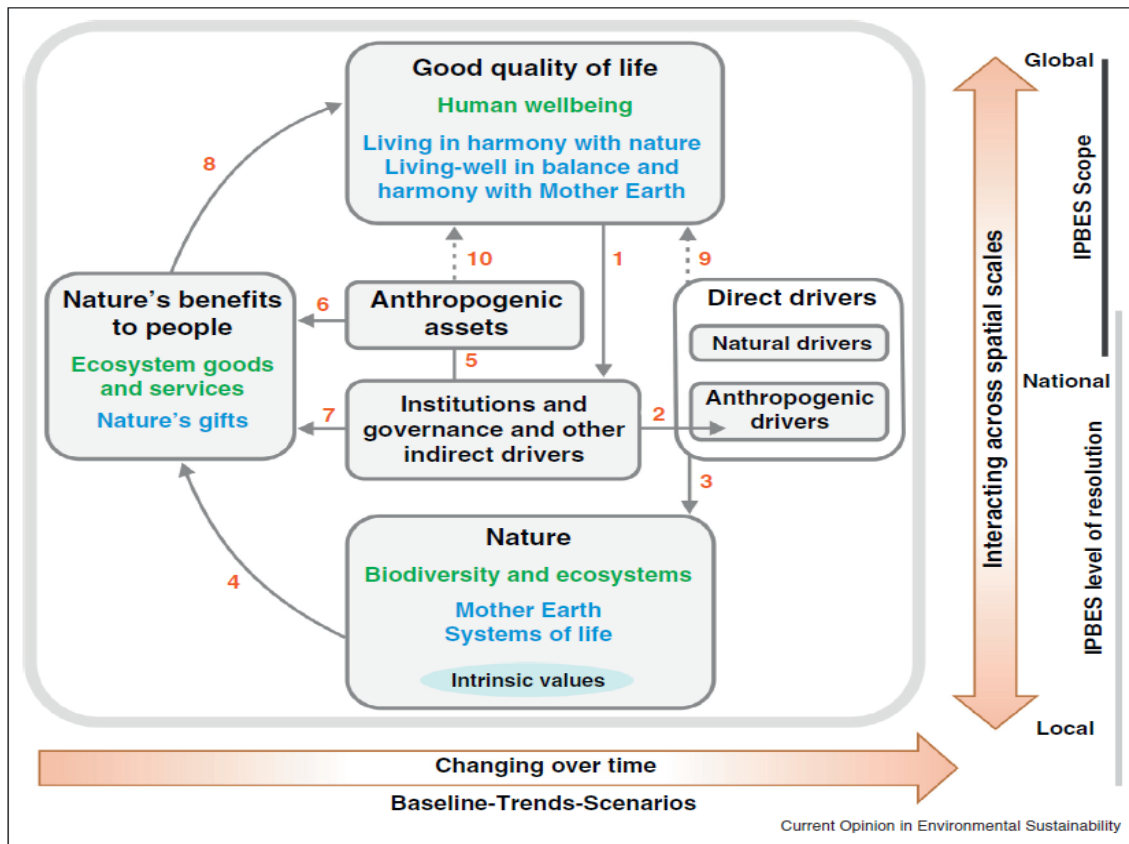


Figura 57. Opinión actual en lo que respecta a la sostenibilidad medioambiental. Fuente: Díaz *et al.*, 2015.

De los fracasos anteriores, es importante extraer aprendizajes, conclusiones (Perrings *et al.*, 2011). El IPBES debía reforzar la relevancia política, por una parte, conectando de una forma mucho más eficiente la evaluación técnica y la gobernanza. De igual manera, tenía que mejorar la evaluación a la hora de tomar las decisiones.

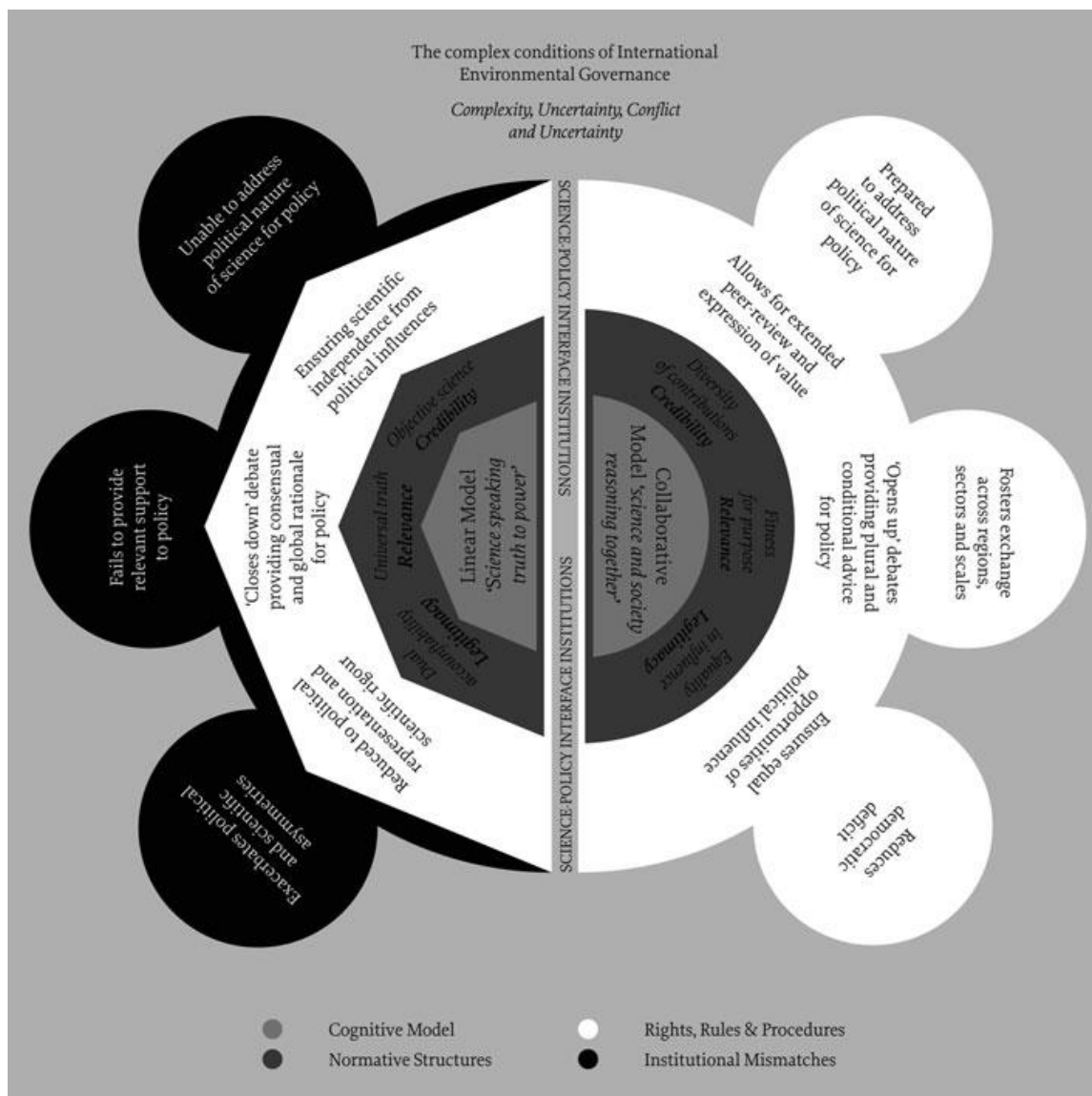


Figura 58. Desajuste y coincidencia de instituciones de interfaz ciencia-política colaborativas frente a las lineales. Fuente: Koetz y Bridgewater, 2012.

El IPBES era una institución emergente y experta asesora cuya finalidad era abordar la pérdida de biodiversidad y la degradación de los servicios de los ecosistemas (Borie y Hulme, 2015). Arrancando del ejemplo del IPCC y de otras experiencias anteriores en materia de biodiversidad, aspiraba a crear un nuevo tipo de interfaz científico-política que lograra alcanzar un equilibrio entre los países desarrollados y los países en vías de desarrollo, incluyendo las diferentes disciplinas y sistemas de conocimiento. Pretendía, también, dar respuesta a la falta de transparencia existente en la gobernanza global de biodiversidad (Turnhout *et al.*, 2014), aunque conseguir un nivel de transparencia no era suficiente; también era necesario estandarizar y gestionar el conocimiento.

3.3. Los retos del IPBES

Los expertos pusieron de manifiesto la necesidad de cubrir el vacío de información científica sobre biodiversidad y servicios de los ecosistemas, haciéndola accesible y sintetizándola, mejorando de esta manera la información puesta a disposición de los responsables de la toma de decisiones. Es de esperar que el IPBES potencie la relevancia de las políticas de biodiversidad llevando a cabo una integración mucho más profunda de los distintos elementos de las políticas científicas (Perrings *et al.*, 2011).

Aunque ya existen muchas organizaciones e iniciativas que ofrecen esa información, tienen limitaciones, y no existe un mecanismo mundial reconocido por instituciones y científicos que reúna y sintetice esa información para la adopción de decisiones en foros normativos como el CDB. Los sistemas y procesos para recoger el conocimiento, servir de base a los que realizan las normas y promover activos diálogos entre ciencia y política en lo que respecta a las cuestiones de la conservación todavía no han dado una respuesta en forma de conocimiento creíble para informar la toma de decisiones de los que encargados de hacerlo (Pe'er *et al.*, 2013).

Por consiguiente, la expectativa es que el IPBES se convierta en el pilar de un cambio radical en lo que respecta a las interfaces globales entre ciencia de la conservación y política. Para que sea creíble, el proceso de elaboración de la información tiene que quedar sujeto a un riguroso proceso de transparencia y revisión de pares (Pe'er *et al.*, 2013). El desarrollo de políticas relacionadas con la conservación de la biodiversidad y los servicios de los ecosistemas, el mantenimiento y uso, requiere que haya disponible un conocimiento científico creíble, relevante y adecuado en términos de tiempo (Nesshöver *et al.*, 2016).

Los científicos señalan que uno de los factores determinantes de que no se haya logrado frenar la pérdida de biodiversidad es que la toma de decisiones políticas no ha tenido en cuenta el conocimiento científico. El CDB, aunque tenía el SBSTTA —teóricamente responsable de dar la visión científica—, no fue capaz de conseguirlo.

Sin embargo, todavía más clave ha sido la falta de indicadores que se comentó más arriba. De hecho, Turnhout *et al.* (2014) señalaron que el IPBES tenía que servir de marco para discutir el papel de la capacidad de medir en la gobernanza de la biodiversidad global. El objetivo del IPBES era garantizar que el conocimiento científico más fiable y actualizado guiase las políticas de biodiversidad, aunque es evidente que para ello hace falta voluntad política. De alguna manera, el IPBES vino a cubrir el vacío que dejó el mal funcionamiento del SBSTTA. Además, el IPBES

tiene también como finalidad promover la creación de indicadores que permitan medir la pérdida de la biodiversidad y de los ecosistemas.

El IPBES ha dedicado mucho tiempo y esfuerzo a alinearse con una compleja estructura de gobernanza que incluye muchos acuerdos globales y que tiene en cuenta la necesidad de afrontar condiciones concretas de carácter regional o local de biodiversidad y pérdida de servicios de los ecosistemas (Beck *et al.*, 2014).

El potencial del IPBES es ofrecer a la ciencia de la conservación la visibilidad que tiene el IPCC. Además, el IPBES ofrece una información científica rigurosa que puede servir de directriz para el compromiso público y el diálogo político (Redford *et al.*, 2012).

De cualquier modo, el IPBES puede jugar un importante papel de comunicación para transmitir la importancia de la biodiversidad y la necesidad de conservarla, aumentando la concienciación ciudadana como ha hecho el IPCC con sus informes periódicos sobre cambio climático.

Como señala Vohland (2011), tiene fundamentalmente tres funciones. La primera es la identificación de los puntos a incluir en la agenda y los diagnósticos; la segunda, la organización del proceso de diagnóstico y, la tercera, ver cómo elaborar observaciones científicamente relevantes. Con la finalidad de conseguirlo, propone que los científicos actúen como personas que alerten de forma temprana, que identifiquen los tópicos pertinentes, que unifiquen a los distintos agentes y que reflejen las características de las distintas regiones y escalas.

Uno de los retos de los que hacen las políticas y tienen que lidiar con cuestiones relacionadas con el medioambiente es la difícil tarea de ponderar las distintas contrapartidas y sus distintas valoraciones por parte de distintas secciones de su electorado, así como tratar de predecir los resultados que tendrán las elecciones que realicen (Spierenburg, 2012).

El IPBES, por consiguiente, es una institución con conocimiento que se posiciona como una interfaz científico-política. Los debates que está generando y las dificultades con las que se está encontrando a la hora de construirse conceptualmente (Borie y Hulme, 2015) ilustran las dificultades que existen para reconciliar dentro de un marco único «todas las maneras disciplinadas de conocer la naturaleza, así como de conceptualizar las relaciones entre el ser humano y la naturaleza» (Jasanoff y Martello, 2004).

Desde la creación del IPBES ha habido tres sesiones. La primera sesión de la asamblea se celebró en Nairobi (PNUMA, 2011), la segunda en Panamá (PNUMA, 2012) y la tercera en Bonn (PNUMA, 2015).

De los resultados de las sesiones celebradas, aún se desprende una falta de consenso en lo que respecta a la interpretación de los servicios de los ecosistemas y en el enfoque necesario. Incluso existe desacuerdo y falta de consenso de naturaleza conceptual del IPBES y en la forma de las relaciones científico-políticas que el IPBES apoyará (Borie y Hulme, 2015).

Sin lugar a dudas, el objetivo del IPBES es lograr el mayor consenso posible, tal vez como respuesta al fracaso que ha tenido el CDB por falta del mismo. Ahora se está trabajando con el IPBES para que tome conciencia de la importancia de la bioinformática y de lo que le pueden ayudar organizaciones como GEO BON o GBIF, que se verán después en el punto relacionado con esta materia.

Todavía a finales de 2017, se sigue poniendo de manifiesto la necesidad para el IPBES de que exista un compromiso serio de los investigadores de las ciencias sociales y humanidades. Teniendo en cuenta que el hombre es parte de la diversidad biológica del planeta, sus comportamientos son clave en los cambios en la biodiversidad (Stenseke. M y Lariguaderie, A. 2017).

3.4. Las diferencias y similitudes entre IPBES e IPCC

Desde su creación en 1988, el IPCC ha sido una referencia imprescindible para conocer la realidad sobre el cambio climático, una plataforma que ha permitido que los conocimientos científicos disponibles fuesen accesibles para todo el mundo — tanto Gobiernos como ciudadanos— y un ejemplo de lo que podría ser el IPBES. De hecho, el objetivo fundacional de ambas plataformas fue el mismo: asegurar que los Gobiernos y otros responsables de tomar decisiones tuvieran a su disposición los mejores conocimientos científicos para poderlos utilizar.

Aunque también existen diferencias, por ejemplo que, mientras las evaluaciones del IPCC se hacían a escala global, IPBES se centraba en el aspecto local y las diferencias regionales, debido a la importancia de gestionar la biodiversidad a escala más reducida. Entre las diferencias existentes, estaba también el hecho de que los encargados de elaborar políticas prestasen más atención a las recomendaciones científicas, pues en el caso del cambio climático, había una diferencia abismal entre lo que pedía el IPCC en sus informes periódicos y lo que se acordó en las cumbres del clima.

No obstante, existen otras organizaciones o instrumentos que conviven y que, en algunos casos, han contribuido al correcto funcionamiento del IPBES. Eso sí, el IPBES tuvo que enfrentar unos retos parecidos al IPCC (Beck *et al.*, 2014). En primer lugar, tuvo que afrontar cómo desarrollar su relevancia política. En segundo lugar, hizo frente la organización y la integración. Tal como se puso de manifiesto, lo afrontó de forma más regional y estableció una multiplicidad de fuentes y formas

de conocimiento y experiencias que le ayudaron a resolver las cuestiones que quería resolver. El tercer reto al que se enfrentó estaba relacionado con la forma de participar en el IPBES y la *accountability*.

La negociación del IPBES dejó ver, una vez más, el intento de los Gobiernos de crear una herramienta científica y, al mismo tiempo, de tratar de influir en la misma, negándose a delegar su soberanía, su control. Algunos científicos, a la vista de la creación del IPBES, manifiestan su expectativa de que pueda servir de puente entre las distintas organizaciones.

En cualquier caso, el hecho de que el IPCC haya sido un éxito en la gestión no quiere decir que pueda ni que tenga que replicarse (Beck *et al.*, 2014). Tal vez la estructura que funcionó tan bien en el caso del cambio climático no sea de aplicación a las cuestiones de biodiversidad. No necesariamente puede replicarse el modelo en todos los casos.

La gran diferencia entre el sistema de lucha contra el cambio climático y el de la lucha para la conservación y el uso sostenible era que el marco científico en el cual se llevaba a cabo el análisis de los problemas relacionados con el cambio climático era mucho más lineal y, por lo tanto, más sencillo de abordar que el de la pérdida de biodiversidad.

En el caso del cambio climático, la complejidad del consenso residía en la negociación de soluciones que permitieran abordar los problemas derivados del análisis científico al estar en juego el modelo económico productivo en su casi totalidad. De hecho, la afectación a los modelos económicos está haciendo que el consenso alrededor del cambio climático esté empezando a debilitarse, a cuestionarse.

Mientras, la problemática en el caso de lucha para la conservación y el uso sostenible radicaba en la práctica inexistencia de sistemas científicamente consensuados para analizar la pérdida de biodiversidad misma (si se produce o no y a qué escala espaciotemporal) y todas las implicaciones que ello tenía. El problema planteado es mucho más complejo, puesto que el rango de datos que se abordan era muy amplio: desde el nivel genético hasta el nivel espacial, del paisaje ecológico o incluso del conjunto del planeta.

Como se verá después, será el GEO BON (por las siglas en inglés de Group on Earth Observation for Biodiversity) el que tome la iniciativa en la creación de indicadores. Empieza a trabajar en las llamadas variables esenciales de biodiversidad, cuya elaboración tiene origen, precisamente, en la falta de linealidad y en la complejidad de la biodiversidad.

Bloque II. La aplicación de las nuevas tecnologías a la biodiversidad: el nacimiento de la informática aplicada a la biodiversidad o bioinformática

Capítulo 1. El nacimiento del término bioinformática

El término como tal lo acuñó John Whitting y representa la aplicación de la tecnología a los datos de biodiversidad. Por lo tanto, se define como la ciencia (Bioinformática, s. f.) que estudia la aplicación de tecnologías computacionales a la gestión y análisis de datos biológicos. Los términos *bioinformática*, *biología computacional* y, en ocasiones, *biocomputación* son utilizados en muchas situaciones como sinónimos y hacen referencia a campos de estudios interdisciplinarios muy vinculados que requieren el uso o el desarrollo de diferentes técnicas estudiadas universitariamente en la ingeniería informática como ciencia aplicada de la disciplina informática. Entre estas pueden destacarse las siguientes: matemática aplicada, estadística, ciencias de la computación, inteligencia artificial, química y bioquímica con las que el ingeniero informático soluciona problemas al analizar datos o simular sistemas o mecanismos, todos ellos de índole biológica y, usualmente (pero no de forma exclusiva), en el nivel molecular.

El núcleo principal de estas técnicas se encontraba en la utilización de recursos computacionales para solucionar o investigar problemas sobre escalas de tal magnitud que sobrepasan el discernimiento humano. La investigación en biología computacional se solapa a menudo con la biología de sistemas.

El primero que usó el término como tal fue Peter Schalk en un artículo de 1998, en el que puso de relieve la distancia existente entre los científicos y los que redactan las políticas. Parecía existir una brecha infranqueable a la que ahora más que nunca hay que poner fin. Era necesario que hubiese una próxima colaboración entre ambos colectivos. En este momento, la creación de datos a gran escala permite que los científicos combinen los conjuntos de datos y creen nuevas herramientas con mayores usos de estos datos.

La bioinformática es un término amplio que incluye la aplicación de la tecnología (informática) a todas las ciencias. La bioinformática es, por lo tanto, la aplicación de las técnicas de ciencia computacional y estadística a la gestión y análisis de los datos biológicos (Jones *et al.*, 2006). Aunque dentro del concepto de biodiversidad, están también las cuestiones más concretas de biomedicina, el objeto de esta tesis, sin

embargo, es más limitado y afecta a la tecnología aplicada al ámbito de la biodiversidad. La bioinformática incluye la aplicación de las tecnologías de la información a la gestión, exploración con algoritmos, análisis e interpretación de los datos primarios relacionados con la vida, en particular en lo que afecta al nivel de organización de las especies (Soberón y Peterson, 2004).

Es un campo que está naciendo y en el que se aplican las herramientas de gestión de la información y análisis a especies, taxonomías y datos recogidos en imágenes (Johnson, 2007). Persigue desarrollar métodos de conocimiento a nivel de organización que lleguen hasta el principio de la historia natural archivada. Su objetivo fundamental es desarrollar hipótesis que cubran la totalidad del árbol de la vida. Nació como una disciplina que unía la información biológica de fuentes contemporáneas e históricas dentro de espectro de la vida (Sarkar, 2007).

Aunque pueda parecer que la bioinformática es un ámbito para los informáticos, esta nueva disciplina requiere de profundos conocimientos de la complejidad de los datos de biodiversidad, que es un conocimiento más propio de los biólogos (Martellos y Attorre, 2012). Hasta la fecha, las mayores tendencias de la bioinformática, la disponibilidad de los datos y de la tecnología han conducido muchas de las ideas y conceptos en el terreno (Peterson *et al.*, 2010). El estado deseable en el terreno hace necesario analizar de qué manera ideas y conceptos conducen el desarrollo de la nueva tecnología y de los nuevos recursos de datos.

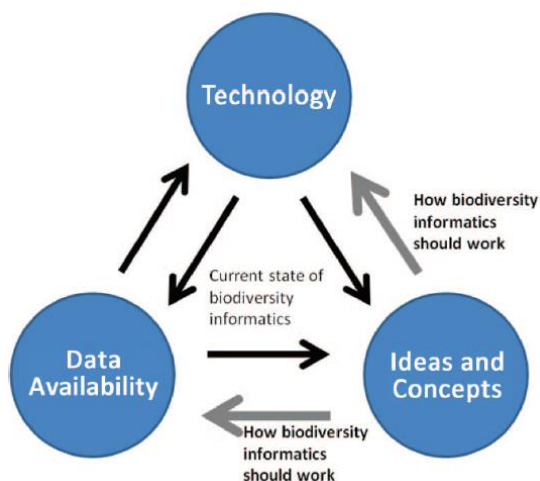


Figura 59. Ilustración de las mayores tendencias en biodiversidad informática: hasta la fecha la disponibilidad de los datos y de la tecnología han sido los motores de muchas de las ideas y de los conceptos en este ámbito. Lo deseable sería que tanto las ideas como los conceptos sean los motores que desarrollen la nueva tecnología, los nuevos datos y los nuevos recursos de datos. Fuente: Peterson *et al.*, 2010.

En la actualidad, la bioinformática gira en torno al concepto de dato y afecta a tres grandes categorías. La primera, la recogida y captura de datos; la segunda, la compilación y servicios de los datos y, la tercera, la visualización de los datos. Los

recientes avances en las tecnologías de la información y un deseo creciente de compartir los datos primarios de biodiversidad están haciendo posible un acceso a la información sin precedentes (Soberón y Peterson, 2004). Peterson analiza lo que realmente podría ser la bioinformática.

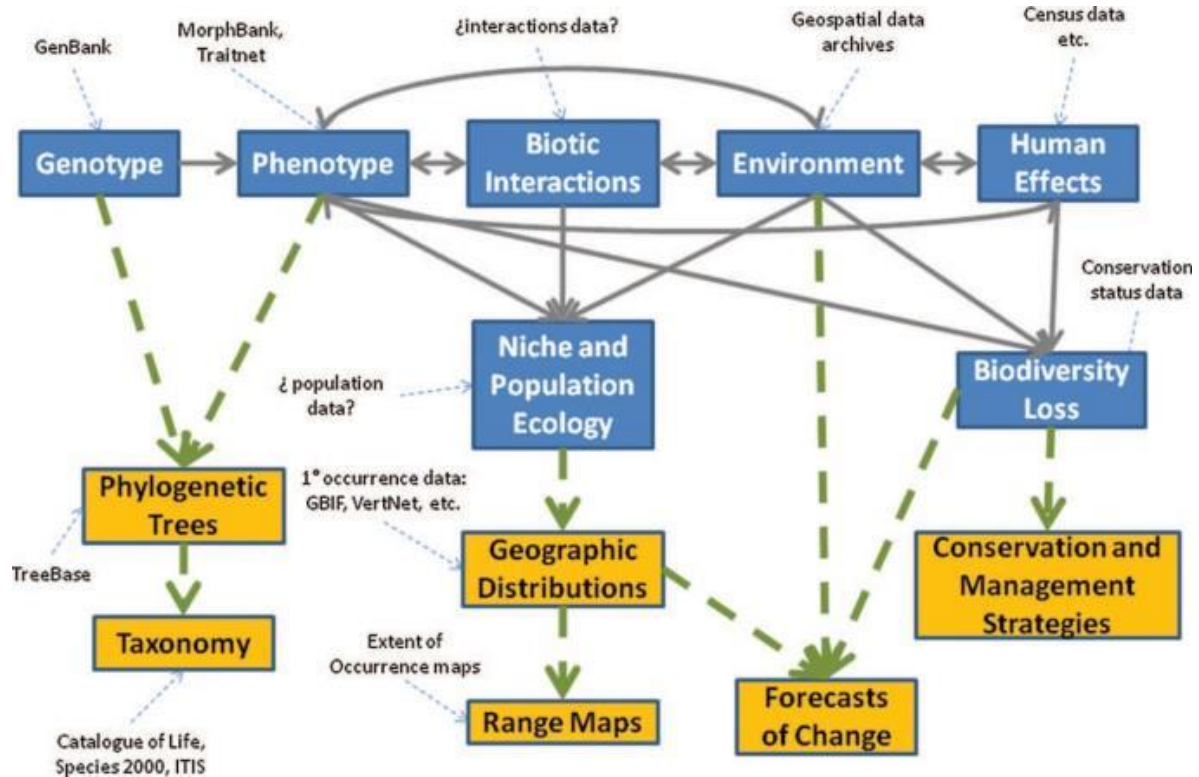


Figura 60. Resumen en los términos más amplios de la biodiversidad informática a nivel mundial. Las cajas azules son los elementos básicos que subyacen en los procesos biológicos, desde el genotipo, fenotipo hasta la ecología y la pérdida de biodiversidad. Las cajas amarillas representan los productos de la información de la biodiversidad que con frecuencia se exploran. Las etiquetas que se encuentran fuera de las cajas son ejemplos de recursos informativos o iniciativas para la mayor parte de elementos del diagrama. Fuente: Peterson *et al.*, 2010.

En el gráfico anterior se muestra un resumen, en términos amplios, del mundo de la bioinformática. Las cajas azules son los procesos biológicos básicos partiendo del genotipo y fenotipo hasta llegar a la ecología y pérdida de la biodiversidad. Las cajas naranjas son los productos de la información sobre biodiversidad.

El reto fundamental que afrontó la bioinformática fue la integración de la información digital procedente de distintas fuentes (Page, 2008). No solo era necesario hacer frente a una disparidad y un sinnúmero de proveedores, cada uno con sus propias comunidades de usuarios, sino que, además, estaban interesados en información distinta (especies de los museos, taxonomías, publicaciones científicas, etc.). La integración requería el uso de identificadores que permitieran poner en relación la información de las distintas bases de datos.

Otro de los retos importantes, como no podía ser de otra manera, era la digitalización del conocimiento existente ligado al hecho de encontrar los fondos necesarios para poder llevarla a cabo (Vollmar *et al.*, 2010). Por eso, han surgido movimientos, que se analizarán con más detalle, para impulsar toda esta digitalización como puede ser el GBIF. Otro ejemplo de digitalización, al haberse detectado una necesidad imperiosa de movilizar datos entre localizaciones y taxonomías, es Vernet, una base de datos de vertebrados (Constable *et al.*, 2010).

La bioinformática ha tenido y tiene un papel fundamental en lo que respecta a los esfuerzos de investigación de la comunidad científica en relación con la sostenibilidad y la conservación (Hardisty *et al.*, 2013). Es necesario entender que la biodiversidad es mucho más que las observaciones de las especies y las taxonomías. Como señalan Canhos *et al.* (2004), aunque la bioinformática es un campo con mucho potencial, todavía hay mucho por hacer. Esta afirmación de 2004 tiene, todavía hoy, plena vigencia. Falta un cierto planteamiento teórico de lo que es la bioinformática en su conjunto que ha llevado a todo un estudio previo.

La bioinformática un campo de conocimiento viejo y joven al mismo tiempo (Peterson y Soberón, 2015). Sus fuentes son ya viejas; sin embargo, partiendo de esa «vieja» información, se está construyendo toda una nueva disciplina que parte de las bases de datos para elaborar toda una construcción flexible de acceso a los datos.

Aunque ha sido una disciplina que ha crecido rápido, no se han planteado las cuestiones científicas generales necesarias para servir de base a su construcción. El resultado de ello es que el desarrollo que ha tenido no tiene vínculos con perspectivas auténticas y progreso a futuro.

La bioinformática existe hoy sin objetivos directores que representen fronteras intelectuales y retos. Este vacío, desafortunadamente, hace que esta disciplina carezca de un marco para un pensamiento efectivo (Peterson *et al.*, 2010). Señalan que la bioinformática tiene que ser un vínculo entre las distintas dimensiones de la biología orgánica (genómica, ecología, etc.).

Por eso, el gran reto la bioinformática es el desarrollo de una metainfraestructura que permita que los datos disponibles puedan utilizarse de forma conjunta con la finalidad de conseguir la modelización y poder hacer frente a las cuestiones relacionadas con el uso del entorno natural capturando la variedad, la diferencia y la complejidad de toda la vida de la Tierra por parte del hombre.

La bioinformática, según Peterson y Soberón (2015) es un ámbito que está en un proceso de rápido crecimiento en e-infraestructuras de datos, herramientas y participación de los investigadores a nivel mundial de distintas disciplinas y con diversos enfoques innovadores. Prosiguen estos autores en el análisis que publicaron en 2015 analizando la bioinformática desde el punto de vista de la evolución en la

última década, poniendo de relieve que su objetivo es abordar fundamentalmente cinco objetivos globales:

1. Movilizar el conocimiento actual.
2. Compartir dicho conocimiento y la experiencia del sinfín de implantaciones a nivel global.
3. Evitar la creación de silos y que se reinventen las herramientas de implantación del conocimiento.
4. Afrontar los retos que tiene adecuadamente.
5. Buscar soluciones a retos difíciles de índole estratégica.

Sin lugar a dudas, la bioinformática se enfrenta a grandes retos, a pesar de los importantes avances que han tenido lugar como la disponibilidad de los datos: el que no sean perfectamente enlazables, reconocibles y accesibles como debieran, etc. (Hardisty *et al.*, 2013). En esta misma línea, se hace necesario construir formas que permitan la creación de un único mapa global de biodiversidad de fácil acceso (Guralnick y Hill, 2009).

La bioinformática es una condición *sine qua non* para que pueda existir este enfoque emergente a la ciencia denominada intensiva en datos (Hey *et al.*, 2009). Hoy, con este enfoque intensivo en datos, lo que se persigue no es solo la consecución de datos en sí misma, sino el desarrollo del conocimiento, mediante el uso de todos los datos que están, hoy, a disposición de los científicos.

Clave en todo este proceso fue la creación, por parte de un conjunto de países entre los que se encontraban, como no podía ser de otra manera, Estados Unidos y la Unión Europea, de la Infraestructura Mundial de Información en Biodiversidad (GBIF, por las siglas en inglés Global Biodiversity Information Facility).

Capítulo 2. Los primeros pasos de la bioinformática: la creación de la Infraestructura Mundial de Información en Biodiversidad

1. LA CONFIGURACIÓN DEL GBIF

Esta organización se configuró como una organización intergubernamental, fue formalmente constituida en 2001 y está integrada, en la actualidad, por cincuenta y tres países y cuarenta y tres organizaciones internacionales.

Se estructuró como una red de nodos nacionales con una secretaría internacional en Copenhague que les daba soporte. El objetivo del GBIF era dar acceso —vía Internet, de manera libre y gratuita— a los datos de biodiversidad de todo el mundo para apoyar la investigación científica, fomentar la conservación biológica y favorecer el desarrollo sostenible.

La visión del GBIF fue lograr un mundo en el que la información relativa a la biodiversidad estuviera libre y universalmente disponible para la ciencia, la sociedad y un futuro sostenible (mecanismo para sacar datos y hacerlos accesibles, digitalización de bibliotecas) (GBIF Secretariat Universitetsparken, s. f.).

2. LA CREACIÓN DEL GBIF

Nació en 1996 a través de un grupo de trabajo de la OCDE (Organización para la Cooperación y Desarrollo Económico) denominado Mega Science Forum Working Group (OECD, 1999) cuyo principal interés era lanzar iniciativas científicas de interés fundamental que por escala no podía abordar ningún país de forma individual. Dentro de este foro, surgió el concepto del GBIF con la idea *de aplicar la informática como mecanismo para facilitar y administrar información proveniente de la naturaleza*.

En octubre de 2010, en la pasada 17.^a reunión de la Junta de Gobierno del GBIF, celebrada en Suwon (República de Corea) el futuro del GBIF quedó asegurado mediante la aprobación de un nuevo Memorando de Entendimiento (conocido

como MOU por las siglas en inglés de Memo of Understanding) que estableció la continuidad de la red de manera indefinida. Los delegados del GBIF aceptaron también un presupuesto para la organización internacional desde 2010 hasta 2016. Además, las delegaciones del GBIF aprobaron por unanimidad la Declaración de Suwon, que subraya la importancia de la red tanto para los Gobiernos como para la ciencia, la conservación y la sostenibilidad.

Las prioridades del GBIF se concentraban en el nivel de organismos y, dentro de este nivel, la prioridad inicial fueron las colecciones. En la actualidad, los cuatro programas que el GBIF tiene en marcha son:

- Tecnología e interoperabilidad (*Data Access and Data Interoperability, DADI*).
- Catálogo de nombres (*Electronic Catalog of Names of Known Organisms, ECAT*).
- Informatización de colecciones de historia natural (*Digitization of Natural History Collection Data, DIGIT*).
- Formación y cooperación (*Outreach and Capacity Building, OCB*).

El GBIF fue concebido como una red de bases de datos interconectadas que pretendía ser una herramienta básica para el desarrollo científico de los países y contribuir significativamente a una mejor protección y uso de la biodiversidad en el planeta. Si no existiera el GBIF, sería necesario crear una infraestructura global que pudiera publicar e intercambiar datos de biodiversidad. A día de hoy se constituye como muestra el siguiente gráfico:

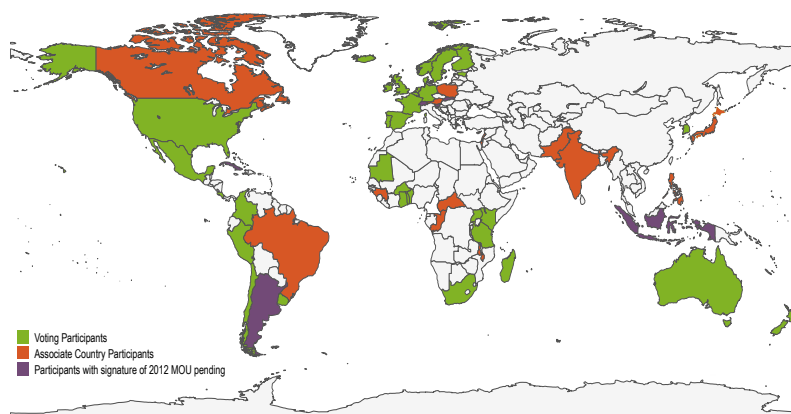


Figura 61. Representación de los países. Fuente: GBIF⁶².

⁶² Recuperado de: www.gbif.org/participation/summary. (Consultado por última vez el 30 de marzo de 2018).

Consta de un órgano de gobierno en el que están representados todos los países miembros, países asociados y organismos institucionales asociados. Un Secretariado Internacional, con base en Copenhague, donde se coordinan las actividades del GBIF y desde donde se proporciona apoyo a los nodos participantes, dado que la Secretaría ha desarrollado una estructura en nodos.

Los nodos participantes están establecidos en los países miembros del GBIF y son los encargados de realizar el grueso del trabajo para hacer realidad los objetivos del GBIF.

Los estados miembros se afilian al GBIF mediante la firma del llamado el Memorando de Entendimiento, comprometiéndose a montar un nodo participante, que, entre otras cosas, debe ser un portal para la información sobre biodiversidad de su incumbencia. Proporciona el mayor punto de acceso a datos de biodiversidad. Desde 2004 facilita el punto de acceso único a datos de especies, desde bases de datos de investigaciones biológicas y colecciones. Gracias a ella los biólogos disponen de un acceso rápido a más de ciento veinte millones de observaciones para usar en muchos análisis biológicos (Yesson *et al.*, 2007).

3. LA MISIÓN DEL GBIF

La finalidad del GBIF, como señala Hobern (2003) no era otra que conseguir que los datos de biodiversidad pudieran circular libremente de forma universal, desarrollando para ello una red internacional de bases de datos con acceso web a los datos primarios de los organismos existentes en el mundo. Esta red debía ser lo bastante flexible para gestionar las necesidades tanto de los proveedores de datos como de los usuarios que quisieran acceder a los datos de biodiversidad.

El GBIF estaba sujeto a los principios rectores en propiedad intelectual (GBIF, 2010):⁶³

- El primero, como no podía ser de otra manera, era el respeto de las leyes aplicables que sirven de marco dentro del cual tiene que operar.
- El segundo se refería al acceso a los datos, puesto que estaba concebido como una infraestructura de libre acceso, en el sentido más amplio del término.
- El tercero, referido a los derechos de propiedad intelectual sobre datos de biodiversidad en la medida en que debía contribuir a la difusión gratuita de los datos sobre biodiversidad, fomentando que los datos de

⁶³ Disponible en: http://www.gbif.es/gbif/ficheros/Principios_rectores_en_DPI.pdf. (Consultado por última vez el 20 de julio de 2016).

las organizaciones afiliadas pasasen a ser de dominio público y pudiesen ser no solo accesibles, sino también reutilizables, con la única obligación de hacer una correcta atribución.

- El cuarto hacía referencia al reconocimiento (*attribution*) y a la necesidad de que se respete en cualquier uso posterior.
- El quinto afectaba al acceso a datos específicos, a fin de cuentas, era una restricción de los derechos de los propietarios⁶⁴ de los datos a bloquear el acceso a los datos.
- El sexto estaba relacionado con la validez de los datos, condición de uso y acceso al GBIF, que no puede garantizar la validez de los mismos.
- El séptimo, la legitimidad de la recopilación de datos, requería que el acceso a los datos de biodiversidad fuera conforme con las normativas aplicables.
- El octavo estaba relacionado con los derechos de propiedad intelectual sobre las herramientas de biodiversidad que desarrollase (motores de búsqueda, etc.).
- El noveno y último hacía referencia a las transferencias de tecnología. Los participantes reconocieron que el GBIF debía impulsar la transferencia no exclusiva de las tecnologías a países en desarrollo, en cuanto estas estuvieran disponibles.

Uno de los objetivos fundamentales del GBIF era contribuir a la digitalización de los datos de biodiversidad, partiendo de los datos de especies en países desarrollados, originariamente recogidos en otras partes del mundo, de manera que los datos pudieran compartirse fácilmente con los países de origen. Esto promueve la colaboración científica (Edwards, 2004). El GBIF promovió la adopción de estándares y protocolos que posibilitasen la realización de intercambios de datos de biodiversidad (Canhos *et al.*, 2004).

3.1. La relación entre el GBIF y el IPBES

El GBIF y el IPBES son, sin lugar a dudas, complementarios. El IPBES nació para desarrollar la interfaz científico-política (GBIF, 2012). No obstante, la realización de análisis fundados depende de que la existencia de datos de biodiversidad fiables, a los que se pueda acceder libremente, con formatos estandarizados, que permitan

⁶⁴ Si es que los datos pueden tener «propietarios», puesto que existe una fuerte corriente que cuestiona que los datos puedan tener dueño. Tal vez sería más correcto hablar de titulares.

realizar análisis científicos y comparaciones de conjuntos de datos de distintos tipos, entre países y en el tiempo.

La finalidad del GBIF era permitir que dichos datos se publicasen, se recopilasen y se intercambiasen. Por consiguiente, el GBIF era una herramienta digital esencial para el IPBES. Si el GBIF no existiera, el IPBES necesitaría construir y apoyar la elaboración de estándares globales, de e-infraestructuras y crear capacidad para movilizar, publicar e intercambiar datos de biodiversidad. Sin embargo, gracias a que el GBIF existe, el IPBES puede actuar de forma mucho más operativa más rápidamente y con menos costes, haciendo uso de las e-infraestructuras existentes.

Los estudios realizados y el acceso al conocimiento científico que sirve de presupuesto a la existencia del IPBES no serían posibles sin herramientas digitales.

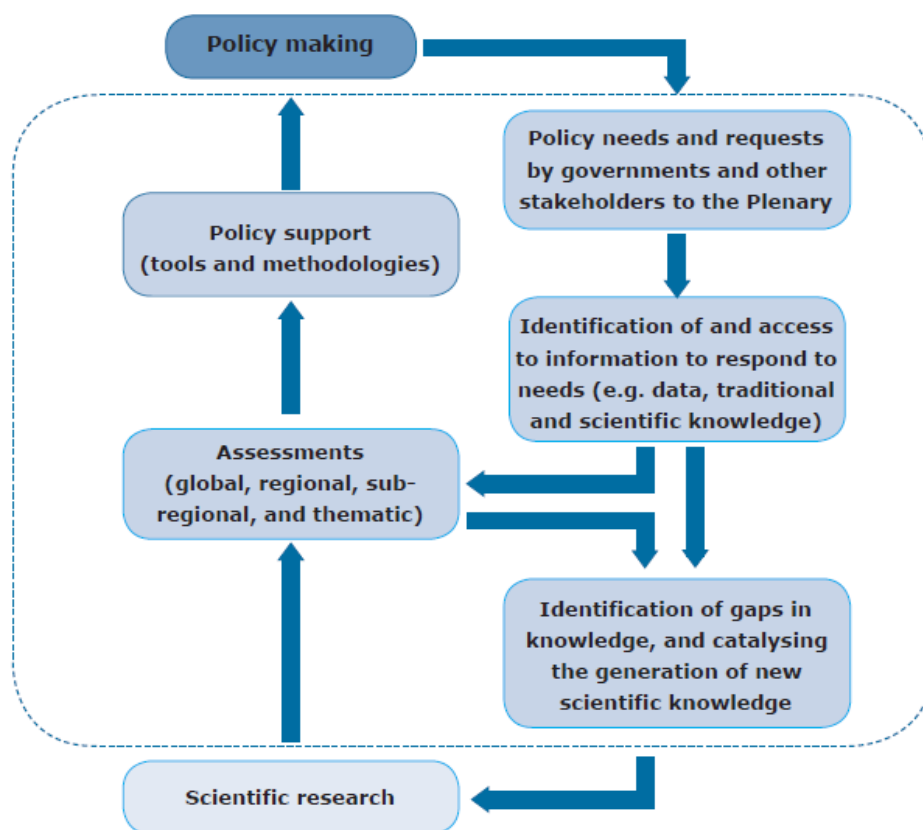


Figura 62. Una posible aproximación a la representación de las funciones del IPBES. Fuente: GBIF, 2012.

El GBIF hizo un análisis de cómo complementa al IPBES poniendo principalmente su atención en los siguientes puntos:

- Los *hallazgos, movilización y acceso a los datos*. El GBIF ha construido la infraestructura de datos para almacenar los descubrimientos que se hagan y permitir el acceso a los mismos.

- *Los estándares de datos, interoperabilidad y principios de acceso abierto.* El GBIF ha desarrollado un amplio rango de herramientas de información y ha trabajado con organismos globales de estandarización para asegurarse de que los datos de biodiversidad publicados dentro de su red cumplen con los estándares internacionales aprobados y cumplen con los requerimientos de acceso abierto. De esta manera, se aseguran de que los datos que tienen no solo son accesibles, sino también interoperables, concepto sobre el que se volverá después, en el apartado relativo a la interoperabilidad.
- *Las e-infraestructuras de información de biodiversidad.* Las movilizaciones de datos del GBIF están organizadas a través de los nodos nacionales, establecidos en instituciones estratégicas.

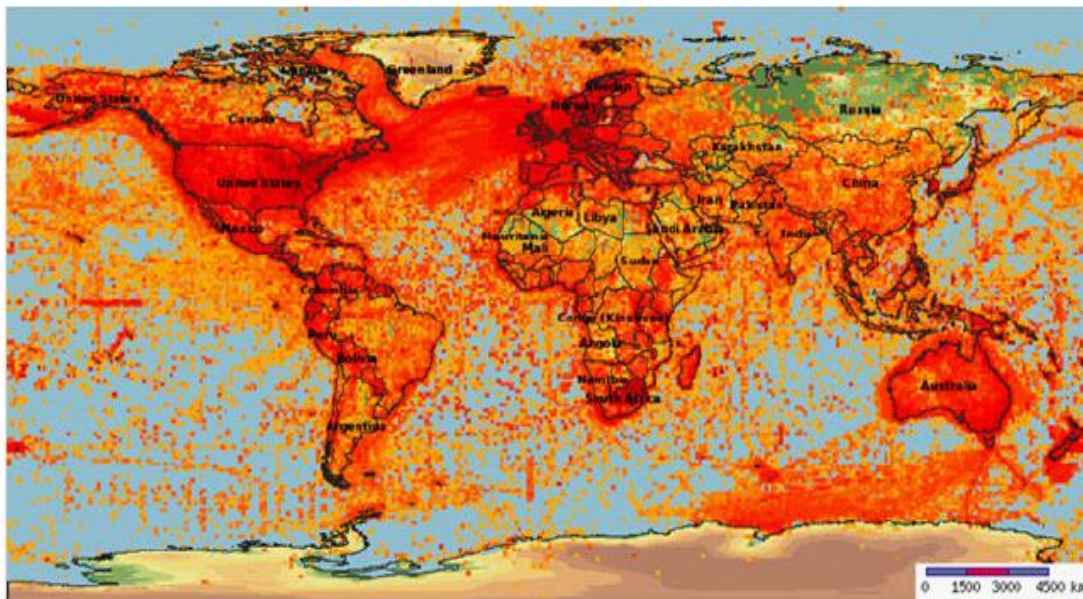


Figura 63. Distribución de los datos a los que se puede acceder a través del portal de datos de GBIF. Fuente: GBIF, 2012.

- *La creación de capacidad.*
- *El uso de datos de biodiversidad para investigación científica.* El portal permite a los investigadores acceder y sintetizar un amplio rango de conjuntos de datos en lo que afecta a las especies.
- El GBIF, además, está *activamente involucrado en otras iniciativas* vinculadas a la biodiversidad como pueden ser el GEO BON, Biodiversity Indicators Partnership, *Conservative Commons*, etc.

Afrontar la crisis de biodiversidad global en la que se encuentra el planeta exige medir la magnitud y la tasa de cambio de forma más efectiva y eso requiere un

acceso fácil y digital a datos de biodiversidad, tanto los históricos como los nuevos, a nivel mundial (Robertson *et al.*, 2013). En el siguiente apartado, se analizará con más detalle el impacto del GBIF como herramienta de la bioinformática.

3.2. El GBIF: una herramienta global de la bioinformática

Algunas organizaciones como el GBIF o el GEOSS ya están usando la bioinformática. A día de hoy se necesitan nuevas herramientas digitales, digitalización para las especies de los museos, etc. (Wheeler, 2004). El GBIF si funciona, y tiene que hacerlo, es una fuente fundamental de conocimiento. Se necesita desesperadamente y cuanto antes, pues es un pilar fundamental del camino que hay que construir.

Algunas personas, con cierta ingenuidad, ven las tecnologías de la información como una forma de liberar los datos de las vitrinas (Wheeler *et al.*, 2004). Sin embargo, la realidad es que, salvo algunas de las taxonomías, la mayoría de los datos están obsoletos o no son fiables.

El rápido acceso a datos «malos» es del todo inaceptable. Hoy el reto no es solo la velocidad de acceso, sino el acceso a datos fiables. Los documentos son dinámicos, están vivos. Ha llegado el momento de construir una ciencia internacional a gran escala. Una de las formas más eficientes de corregir estos errores es hacer visibles los datos (Edwards, 2004) de manera que todos los investigadores puedan comprobarlos y corregirlos.

El GBIF se está consolidando como una importante herramienta de compartir conocimiento, abierta, a la que se puedan afiliar países, organizaciones, etc. La preocupación por poder acceder a los datos apoyando los objetivos de la convención y sus actividades quedó recogida por la secretaría que previó un cierto número de iniciativas de bioinformática y organizaciones con la finalidad de facilitar el acceso a las fuentes por parte de partes y Gobiernos, repatriar información y potenciar capacidades nacionales y regionales en relación con el uso de formatos, protocolos y estándares comunes (Silva, 2004).

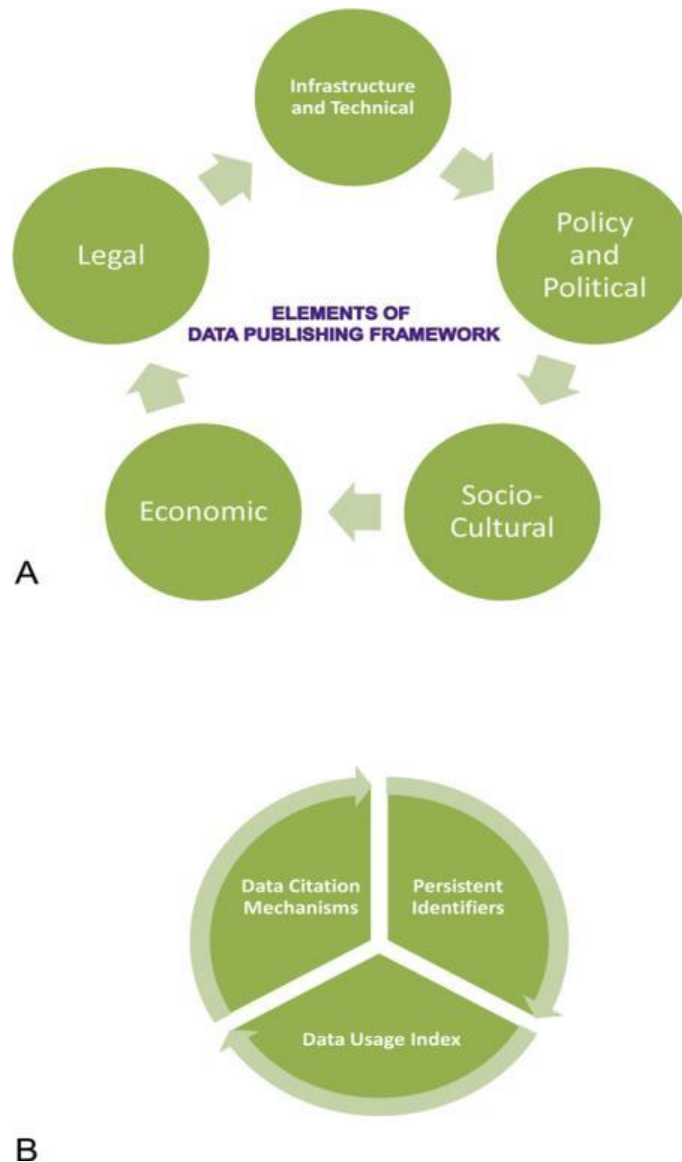


Figura 64. Reflejo de las limitaciones que existen. Fuente: Chavan e Ingwersen, 2009.

Existe una clara necesidad de institucionalizar un marco de publicación de datos (Moritz *et al.*, 2011) que pueda abordar las limitaciones socio-culturales, técnicas y de infraestructuras, normativas y de políticas (Chavan y Ingwersen, 2009), así como hacer frente a las cuestiones relacionadas con la sostenibilidad y apoyo financiero.

Con esta finalidad, un enfoque estandarizado hacia las definiciones formales y publicaciones de los datos en el ámbito de los datos de biodiversidad, el GBIF (el organismo intergubernamental al que más claramente se le puede encargar esta tarea) convocó un Grupo de Trabajo de Marco de Publicación. Se concibió este grupo de trabajo como un entorno conductor para garantizar un acceso gratuito y abierto a los datos mundiales de biodiversidad.

La era digital presenta nuevas oportunidades para la comunidad científica (Research Information Network, 2008). Los hallazgos científicos en formato digital pueden

moverse más fácilmente, duplicarse, enviarse a otros, mezclarse de nuevas maneras, dividirse en nuevas formas, almacenarse en grandes cantidades y manipularse en supercomputadores si su naturaleza así lo requiere. Existe hoy un extendido reconocimiento de que los datos son valiosos a largo plazo y de que compartirlos y ponerlos a disposición del público es fundamental.

Dos razones fundamentales justifican ponerlos a disposición del público: la primera, que formen parte de los registros académicos con la finalidad de que puedan validarse y probarse y, la segunda, que puedan reutilizarse por otros en una nueva investigación. La publicación de los datos no es suficiente, pues es necesario publicar también los metadatos (Roberts y Moritz, 2015).

Si bien es innegable la contribución del GBIF a la proliferación de datos —en 2012 había ya más de 377 millones de registros compartidos—, la bioinformática se centra también en la calidad de los datos. Por eso, el GBIF tiene vocación de convertirse en una forma de filtrar que la información compartida es relevante para el conocimiento científico (Gaijy *et al.*, 2013). Se han identificado ciertos vacíos en esta infraestructura que hacen que se estén llevando a cabo trabajos para mejorar la calidad de los datos.

De hecho, el grupo de trabajo del GBIF ha investigado las necesidades de los usuarios en lo que respecta a los datos de biodiversidad identificando una serie de áreas principales de mejora para movilizar los datos que respondan a estas necesidades. Para ello hace tres tipos de recomendaciones: la primera, relacionada con los vacíos de datos, los volúmenes de datos y la calidad de los mismos; la segunda, relacionada con los nuevos tipos de datos y las nuevas aplicaciones y, la tercera y última, destinada a promover la facilidad de uso y los incentivos para un mayor uso (Faith *et al.*, 2013).

3.3. El GBIF *Data Sharing*

Los datos sobre biodiversidad se encuentran dispersos en muchas bases de datos o se encuentran en papel u otros medios no susceptibles de búsqueda interactiva (Otegui Tellechea, 2012). El GBIF es un marco para facilitar la digitalización de los datos sobre biodiversidad y para hacer interoperable un número aún desconocido de bases de datos sobre diversidad biológica distribuidas en todo el mundo.

En colaboración con otros esfuerzos existentes, el GBIF persigue catalizar la finalización de un catálogo de los nombres de organismos conocidos y desarrollará motores de búsqueda para extraer grandes cantidades de datos de biodiversidad. Tiene vocación de convertirse en una herramienta sobresaliente para los científicos, los administradores de recursos naturales y los encargados de realizar políticas (Edwards *et al.*, 2000).

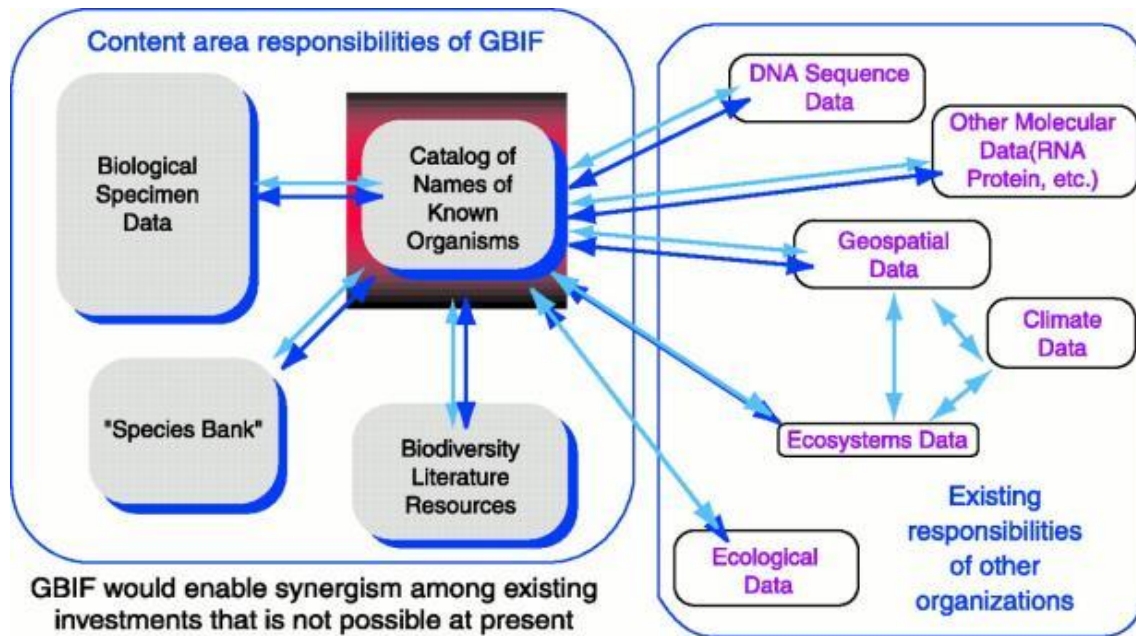


Figura 65. Catálogo electrónico de los nombres de organismos conocidos que servirán de vínculo entre muchos de los tipos de bases de datos biológicos y no biológicos posibles. Esos vínculos permitirán una minería de datos inimaginable hasta ahora porque, en la actualidad, es difícil, si no imposible, descubrir correlaciones entre los distintos conjuntos de datos. Fuente: Edwards *et al.*, 2000.

El GBIF hace posible que se vayan llevando a cabo unas sinergias imposibles hoy, de modo que se logre la interoperabilidad no solo entre bases de datos de biodiversidad. Los datos principales que el GBIF enlaza se denominan datos primarios de biodiversidad (PBD) y son la unidad más básica, libre de interpretación, de los estudios de biodiversidad: qué organismo ha sido observado o muestreado y dónde (Guralnick y Hill, 2009).

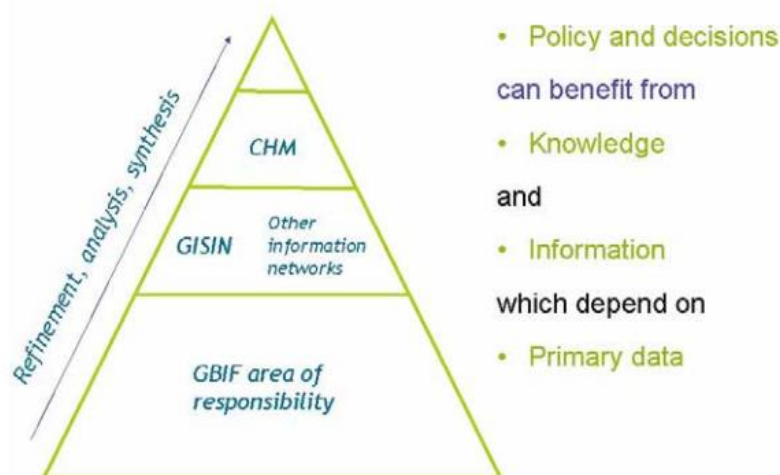


Figura 66. GBIF, en la actualidad, se está centrando en los datos primarios de biodiversidad, en los que se encuentran en la base de la pirámide de información. Fuente: Saarenmaa, 2005.

El GBIF es una herramienta desarrollada con la finalidad de compartir datos (Michener, 2015). Su arquitectura es la que se muestra en el gráfico siguiente (Saarenmaa, 2005).

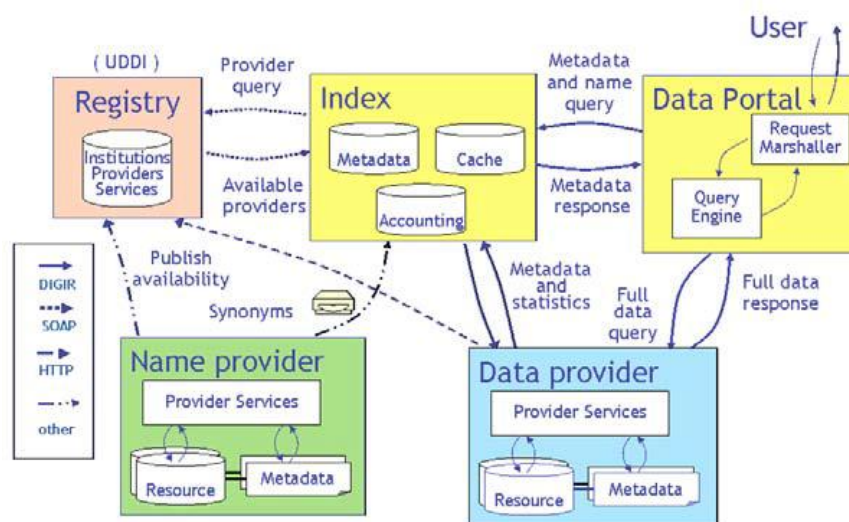


Figura 67. Componentes de la arquitectura del sistema de información de GBIF. Fuente: Saarenmaa, 2005.

El acceso abierto a los datos primarios de biodiversidad es esencial tanto para posibilitar la toma de decisiones como para dar poder a los *stakeholders* involucrados y a los que afecta la conservación de la biodiversidad (Chavan y Ingwersen, 2009).

Existe un gran consenso entre los científicos sobre que los datos deben ser abiertos (Anon., 2003). Por lo tanto, la norma esperada debe ser el intercambio de datos y a estos efectos, Moritz *et al.* (2011) dan hasta 24 recomendaciones.

La primera de las recomendaciones es que todos los datos relevantes para la comprensión y conservación de la biodiversidad deben ponerse a disposición de todos de forma libre, abierta y efectiva. De esa recomendación principal, resultan toda una serie de recomendaciones con la finalidad de que el GBIF establezca una serie de pautas para llevar a cabo lo establecido en la primera recomendación.

La segunda de las recomendaciones es que el GBIF tiene que reexaminar su modelo actual de respaldo de recursos de datos y examinar la práctica actual que los nodos nacionales o los nodos asociados participantes deben endosar antes de que los datos sean descubiertos e indexados a través de la red GBIF.

La recomendación 19 establece que el GBIF debería fomentar que los espónsores de la investigación en biodiversidad, con independencia de que se trate de agencias gubernamentales, corporaciones o fundaciones privadas, establezcan requisitos obligatorios para el *open access* y libre a los datos de biodiversidad. El GBIF debe

promover que las negociaciones de los costes indirectos de los financiadores incluyan cálculos de costes para una infraestructura digital sostenida adecuada para el intercambio gratuito y abierto y para el mantenimiento sostenido, seguro y persistente de los datos. Se debe esperar que las propuestas incluyan una planificación adecuada y una provisión financiera para la gestión y el acceso sostenido de datos. Recomendamos además que el GBIF fomente procesos de revisión por pares que incluyan una rigurosa revisión de historias pasadas de intercambio exitoso y que respalde la norma de la planificación de última generación para compartir, no simplemente promesas de «poner datos en la web».

En esta misma línea, se hacen hasta 24 recomendaciones adicionales que no tiene sentido enumerar aquí y que pueden consultarse en el texto mencionado.

4. EL GLOBAL BIODIVERSITY INFORMATICS OUTLOOK (GBIO): DELIVERING BIODIVERSITY KNOWLEDGE IN THE INFORMATION AGE

4.1. El origen del GBIO (GBIF, 2015)

No pueden analizarse los resultados del GBIO sin entender antes qué es. Este documento fue desarrollado mediante una consulta a la comunidad, en un proceso que inició la Global Biodiversity Informatics Conference (GBIC) (Hobern *et al*, 2012).

La conferencia reunió a cerca de cien expertos de una gran variedad de disciplinas que se encontraron en 2012 en Copenhague. Científicos, informáticos, políticos, juristas y otros fueron invitados a identificar la mejor manera de aprovechar el poder de las tecnologías de la información, la ciencia de la biodiversidad y las redes sociales para mejorar nuestra comprensión de la vida en la Tierra. Mediante una serie de talleres, fueron identificadas las cuestiones más prioritarias y las ineficiencias del GBIF, de GEO BON y de otros sistemas que se analizarán más adelante (LifeWatch, Alas, DataONE, etc.), así como las herramientas necesarias para dar respuesta a esto y subrayar los pasos que hay que dar para crear dichas herramientas e implementarlas de forma efectiva.

Las ideas y prioridades identificadas y debatidas durante las discusiones en el GBIC las estructuraron y ordenaron; para terminar, los líderes de los talleres y un grupo más pequeño redactaron el documento final. La finalidad del documento que recoge el GBIO era presentar el enfoque de consenso y un marco para las personas e instituciones que fueron actores clave en su implementación.

4.2. Las conclusiones del GBIO

Las conversaciones y discusiones sobre preservación de la biodiversidad son cada vez más fuertes y frecuentes. Aunque el IPCC consiguió poner centrarse en el cambio climático, hasta el día de hoy parece que ninguna de las iniciativas dirigida a las cuestiones de biodiversidad ha tenido el mismo impacto.

El GBIO (GBIF, 2015) arranca precisamente de que no se ha alcanzado el objetivo que los Gobiernos de todo el mundo se habían marcado para 2010: reducir el ritmo de pérdida de la biodiversidad biológica. Una de las razones fundamentales ha sido que la información disponible no era suficiente. Generar políticas adecuadas para proteger los hábitats requería comprender lo que dichos hábitats contenían, cómo interactuaban las especies dentro de los mismos, de qué manera respondían a los cambios y presiones, naturales y cuáles de estos cambios tenían origen en la intervención del hombre.

En efecto, tal como señaló el GBIO, con la adopción del Plan Estratégico de Biodiversidad 2011-20, donde quedaron incluidos los objetivos AICHI, los Gobiernos pusieron una vez más de manifiesto la importancia de preservar y restaurar la biodiversidad y mantener los servicios de los ecosistemas del planeta. La bioinformática era el medio para la consecución de estos objetivos, no solamente una herramienta.

Si bien en los últimos doscientos cincuenta años se han generado cantidades ingentes de datos ligados a la investigación de la biodiversidad, muchos no son todavía accesibles.

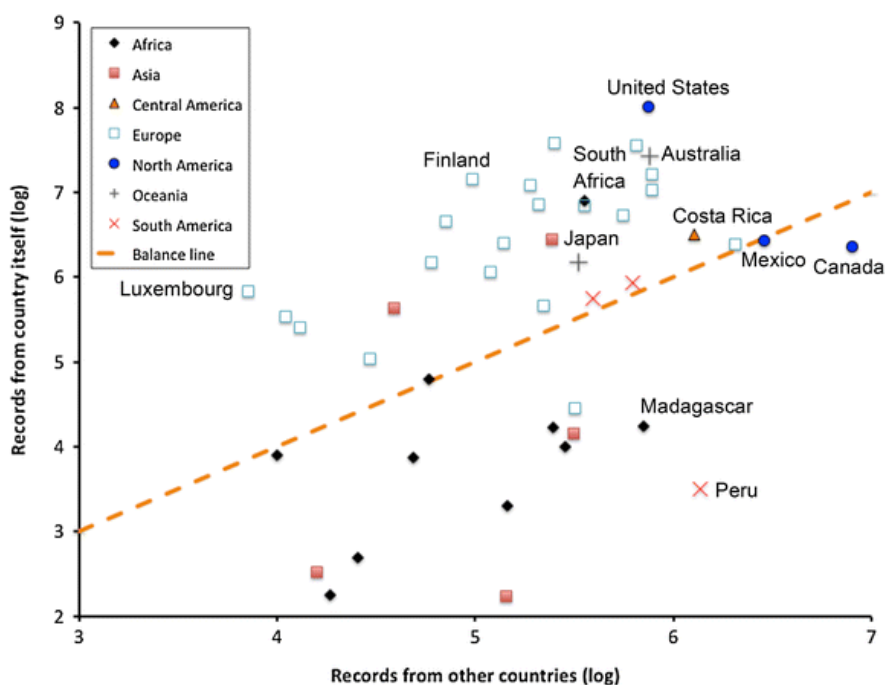


Figura 68. Resumen del conocimiento digitalmente accesible por países mundialmente, dibujado por la Global Biodiversity Information Facility en enero de 2014, que muestra los 10 registros que presentan los países versus lo que proporcionan las instituciones de otros países. Los países (muchos de distintos continentes), que no dan datos, no aparecen en el gráfico. La línea de puntos indica equilibrio entre los datos entregados y los recibidos por el país. Fuente: Peterson y Soberón, 2015.

Sin embargo, hoy, con las nuevas tecnologías, la multiplicación de los datos en la nueva sociedad del conocimiento y de la colaboración a la hora de movilizar los datos de biodiversidad, nuevos y viejos, de forma estructurada y estandarizada, permite un amplio margen de usos, lo que da lugar a nuevas oportunidades de investigación y contribuye al saneamiento de las políticas relacionadas con la biodiversidad.

Pretendía ofrecer un marco para alcanzar un conocimiento mucho más profundo del mundo de la biodiversidad y, a través de ello, comprender lo que significa conservarlo mejor y usarlo de manera más sostenible. Identificó, para ello, cuatro áreas estratégicas, integradas por una serie de elementos, con la finalidad de contribuir a la coordinación de esfuerzos y de financiación:

- *Cultura*: crear una cultura de conocimiento compartido, de estándares robustos y comunes de datos, de políticas e incentivos para lograr su intercambio y un sistema de almacenamiento continuo y de archivo de esos datos.
- *Contenido (datos)*: movilizar los datos de biodiversidad de todas las fuentes existentes para que estén disponibles con rapidez y de forma rutinaria. La recogida o captura de datos solo tiene que producirse una vez, de modo que pueda usarse todas las veces que sea necesario. La referencia a datos incluye cualquier tipo, tanto los de literatura histórica como las colecciones de observaciones realizadas por los ciudadanos científicos, desde la lectura de los sensores automatizados al análisis de las firmas genéticas de las comunidades de microbios.
- *Hallazgos y acceso a datos*: proporcionar las herramientas para convertir los datos en hechos y que sea posible el acceso a los mismos, organizarlos en vistas que los contextualicen y les den sentido. Además, es necesario realizar esfuerzos de colaboración para mejorar la precisión de los datos y su adecuación a investigación y redacción de políticas, con la finalidad de facilitar un marco taxonómico y de organizar la información en relación con las características de las especies y las interacciones entre ellas.
- *Comprensión (modelización y visualización)*: generar una mejor comprensión de la biodiversidad y de cómo impacta en ella el comportamiento del

hombre, mediante la aplicación de modelos y herramientas de visualización e identificando lagunas que permitan priorizar futuras recogidas de datos.

Por consiguiente, como puso de manifiesto el GBIO, la investigación en biodiversidad persigue comprender la variación dentro y entre las especies y su relación geográfica, ecológica, temporal y con los factores antropogénicos. Explora las interacciones entre los organismos y su entorno, incluida la relación con el ser humano y entre los organismos y el entorno. Avanzar en la investigación puede explorar las tendencias, analizar los motores de cambio y realizar pronósticos en lo que respecta al futuro. Todo ello requiere acceso a los mejores datos disponibles registrados en lo que a observación de cada especie se refiere, con el apoyo de la mejor comprensión de los sesgos e incertidumbres asociados a cada conjunto de datos.

La bioinformática es la ciencia que pone en relación el uso de la tecnología con la finalidad de sostener estas necesidades, organizando el conocimiento en relación con los organismos biológicos individuales y los sistemas ecológicos que componen. Con el tiempo, la bioinformática proporcionará una fuente cada vez mayor de interconexión de datos que soporten la investigación científica en el mundo natural.

La primera década del milenio ha sido testigo de un creciente interés en la bioinformática y en la emergencia de una comunidad de científicos y de profesionales de informática, que han colaborado en el desarrollo de nuevas funcionalidades en este ámbito. Existen un cierto número de iniciativas establecidas para llegar más allá en el logro de metas en lo que respecta a la investigación de biodiversidad a través de la informática: Species 2000 Catalogue of Life (CoL),⁶⁵ Biodiversity Information Standards (TDWG),⁶⁶ Global Biodiversity Information Facility (GBIF),⁶⁷ la Encyclopedia of Life (EOL),⁶⁸ el Consortium for the Barcode of Life (CBOL),⁶⁹ la Biodiversity Heritage Library (BHL)⁷⁰ y el Group on Earth

⁶⁵ Información disponible en: <http://www.sp2000.org>. (Consultado por última vez el 29 de mayo de 2016).

⁶⁶ Información disponible en: <http://www.tdwg.org>. (Consultado por última vez el 29 de mayo de 2016).

⁶⁷ Información disponible en: <http://www.gbif.org>. (Consultado por última vez el 29 de mayo de 2016).

⁶⁸ Información disponible en: <http://eol.org>. (Consultado por última vez el 29 de mayo de 2016).

⁶⁹ Información disponible en: <http://www.barcodeoflife.org>. (Consultado por última vez el 29 de mayo de 2016).

⁷⁰ Información disponible en <http://www.biodiversitylibrary.org>. (Consultado por última vez el 29 de mayo de 2016).

Observations Biodiversity Observation Network (GEO BON),⁷¹ entre otras muchas.

A pesar de estas iniciativas, y de muchas otras más tanto en el ámbito nacional como en el regional y en el global, el panorama de la bioinformática continúa estando muy fragmentado. El reto del GBIO era ayudar a coordinar no solamente estos esfuerzos, sino también las contribuciones a toda la investigación sobre biodiversidad.

Se elabora, por lo tanto, un marco que hace hincapié en los puntos prioritarios. Se invita a individuos, organizaciones y naciones a contribuir. En cada uno de los puntos, desarrolla un conjunto de acciones más detalladas a realizar durante los siguientes diez años, con la finalidad de que la visión puesta de manifiesto en el marco del GBIO pueda hacerse realidad.

Organizado en cuatro áreas estratégicas, las divide a su vez en elementos clave cuya consecución evalúa y puntúa con la finalidad de analizar el nivel de consecución de las mismas. Esto se representa poniendo líneas o guiones debajo de cada uno de los elementos, el número de líneas o guiones (entre una y cuatro) refleja el nivel de desarrollo. Las cuatro áreas están interconectadas y se retroalimentan entre sí. Cada una de las áreas y de los elementos individuales que incluye puede desarrollarse de forma independiente y, sin embargo, en la medida en que se desarrollen, se irán retroalimentando y reforzando unas a otras, de modo que todas, conjuntamente, lleguen mucho más lejos que la suma de las partes.

En la base de todo se encuentra el área que se concentra en la cultura. Pone en su sitio los elementos necesarios para que la información sobre la biodiversidad se convierta en un recurso común e interconectado, un almacenamiento estable y duradero, un conocimiento, una cultura y unas políticas que faciliten que se pueda compartir y que existan estándares comunes de datos (GBIF, 2015).

4.2.1. El área de cultura



Figura 69.

Cinco puntos de cultura. Fuente: GBIF, 2015.

⁷¹ Información disponible en: <http://GEO BON.org>. (Consultado por última vez el 29 de mayo de 2016).

Pretende poner las bases necesarias para convertir los datos de biodiversidad en un recurso conectado, compartido de forma abierta y libremente disponible. Los elementos que integran la cultura son las claves esenciales para cualquier otra actividad. Son requerimientos que la bioinformática comparte con otras áreas de investigación orientadas a datos y con otras numerosas e-infraestructuras nacionales que están empezando a gestionar: el almacenamiento estable y duradero, una cultura de uso abierto, estándares de datos adecuados, compromiso con toda la comunidad científica y apoyo de las políticas nacionales e institucionales.

Los elementos clave en el ámbito de la cultura son (Hobern *et al.*, 2012):

- *Transparencia*: la creación de una cultura de acceso abierto y de reutilización de datos. Es necesario estandarizar prácticas de compartir de forma abierta los datos mediante financiación pública y otro tipo de incentivos públicos, y mediante la correcta atribución y reconocimiento de las fuentes primarias de datos, de los creadores de datos y de los que conservan los datos, tanto individuos como instituciones.
- *Estándares de datos*: supone entregar conjuntos estandarizados de datos que permitan la reutilización e interoperabilidad de todos los datos de biodiversidad.
- *Almacenamiento duradero y archivo*: proporcionar una red de repositorios de datos para todo tipo de datos de biodiversidad, con una gestión consistente de metadatos, identificadores, herramientas y servicios.
- *Políticas incentivadoras*: asegurarse de que las políticas públicas, la legislación y las iniciativas de financiación a todos los niveles se combinan para reforzar esta estrategia y para servir de base a los elementos individuales que lo componen.
- *Red de conocimiento de biodiversidad*: crear una estructura técnica que permita la conservación y anotación de los datos usando los mejores conocimientos de la comunidad de manera que haga dicha conservación inmediatamente visible para los futuros usuarios y que proporcione, al mismo tiempo, *feedback* a los tenedores de datos.

4.2.2. El área de datos

Sobre esta base, prosigue el GBIO, el área enfocada en los datos persigue acelerar la movilización de datos de todo tipo de fuentes, desbloqueando el conocimiento de las colecciones y literatura, mejorando la calidad de los datos y subsanando las carencias, así como poniendo en línea las observaciones y los datos de todas las fuentes desde satélites a genomas.



Figura 70. Cinco puntos relacionados de datos. Fuente: GBIF, 2015.

Su finalidad es movilizar los datos de todas las fuentes y organizar estas de manera que puedan soportar análisis y modelizaciones a gran escala (Hobern *et al.*, 2012). Los elementos de datos son propios de la bioinformática. Este área persigue, por consiguiente, movilizar todas las fuentes de conocimiento sobre biodiversidad tanto históricas como actuales, para convertirlas en fuentes de conocimiento accesibles en red y de forma abierta, mediante la digitalización de la literatura y de las colecciones de historia natural y mediante la integración de las observaciones y mediciones de la ciencia ciudadana,⁷² investigación en el terreno, estudios biomoleculares y teleobservación.

Los elementos que integran esta área son:

- *Materiales publicados*: mecanismos para extraer los datos de biodiversidad, en la actualidad, integrados en las publicaciones y otros formatos multimedia con la finalidad de ponerlos a disposición de forma abierta, estandarizada y estructurada.
- *Colecciones y especímenes*: desarrollar y compartir técnicas más eficientes para acelerar los esfuerzos para digitalizar y capturar los datos históricos de las colecciones.
- *Estudios y observaciones de campo*: capturar todas las observaciones sobre biodiversidad, incluyendo sonidos e imágenes, y ponerlos a disposición en cuanto estén realizados, o dentro de un periodo de tiempo determinado.

⁷² «Se entiende por ciencia ciudadana a la investigación científica llevada a cabo por una suma de colaboradores, en su totalidad o en parte por científicos y profesionales junto a gente común. Formalmente, la ciencia ciudadana ha sido definida como “la recopilación y análisis sistemático de datos, el desarrollo de la tecnología, las pruebas de los fenómenos naturales, y la difusión de estas actividades por los investigadores sobre una base principalmente vocacional”». Ciencia ciudadana (s. f.). En Wikipedia. Recuperado de https://es.wikipedia.org/wiki/Ciencia_ciudadana. (Consultado por última vez el 11 de junio de 2016).

- *Secuencias y genomas*: capturar todos los datos relevantes de la actividad genómica, incluyendo las secuencias referenciadas, los metagenómicos medioambientales, la variación genética y la totalidad de los genomas.
- *Observaciones automáticas y a distancia*: explorar las oportunidades para las grabaciones automatizadas y semiautomatizadas y la identificación de especies y poblaciones de fuentes que varían entre las imágenes de satélites hasta las secuencias de genes automatizados.

4.2.3. El área de evidencias

El área enfocada en evidencias afecta a la depuración, estructuración y evaluación de los datos para mejorar la calidad y colocarlos dentro de un marco taxonómico que organice toda la información conocida en relación con las especies (GBIF, 2015).

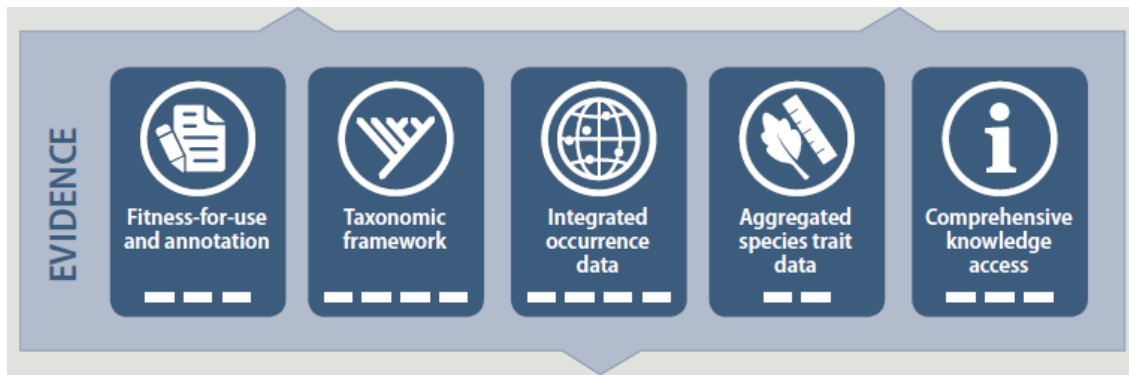


Figura 71. Cinco puntos relacionados con las muestras. Fuente: GBIF, 2015.

Su finalidad es proporcionar las herramientas para dar un apoyo consistente y una indagación comprensiva y global, y un uso de los datos procedentes de todas las fuentes en relación con la biodiversidad de cualquiera de las áreas con el tiempo, de manera que queden cubiertos todos los grupos taxonómicos.

Se centra en las herramientas y servicios requeridos para organizar y mejorar dichos flujos de datos. Muchas actividades nacionales y globales ya están trabajando sobre estas cuestiones, desarrollando los enfoques y herramientas necesarios para potenciar la calidad de los datos y para dar acceso a los elementos clave de los datos de biodiversidad del mundo, en particular en lo que respecta a la información relativa a los nombres y a las clasificaciones, datos de distribución de especies y características de las especies e interacciones.

Los elementos que la integran son:

- *Adecuación al uso y anotaciones*: un mecanismo eficiente que permita a todos (*amateurs* y expertos) gestionar herramientas automáticas y anotar elementos de datos individuales para mejorar su calidad y adecuación,

para usarse con finalidades particulares y para asegurarse de que estas anotaciones se hagan, solamente, una vez.

- *Marco taxonómico:* un catálogo comprensivo, conservado por un experto, con datos de nombres, clasificación y filogenia y que incorporan, también, taxonomías sin nombres formales.
- *Integrar datos de ocurrencias:* poner conjuntamente datos de todas las fuentes para documentar las ocurrencias conocidas de todas las especies, en el tiempo y en el espacio.
- *Características agregadas de datos de especies:* suministrar las herramientas para poner en común todos los datos disponibles de especies e interacciones, y para asegurarse de que se mantiene en formas que permitan su uso en formato digital en análisis y modelizaciones.
- *Acceso comprensivo al conocimiento:* conectar todo el conocimiento publicado en materia de biodiversidad y hacerlo accesible a través de una rica indexación de literatura de biodiversidad, multimedia y otras fuentes, incluyendo la presentación de la información y de las páginas de especies y vía servicios web.

4.2.4. El área de la comprensión

Por último, el área enfocada en la comprensión hace posible realizar una síntesis más amplia, proporcionando las herramientas básicas de modelización con la finalidad de poder echar un vistazo a los ecosistemas en su totalidad, realizar mejores decisiones de política y poder reaccionar a cualquier cambio (GBIF, 2015; Hobern *et al.*, 2012).

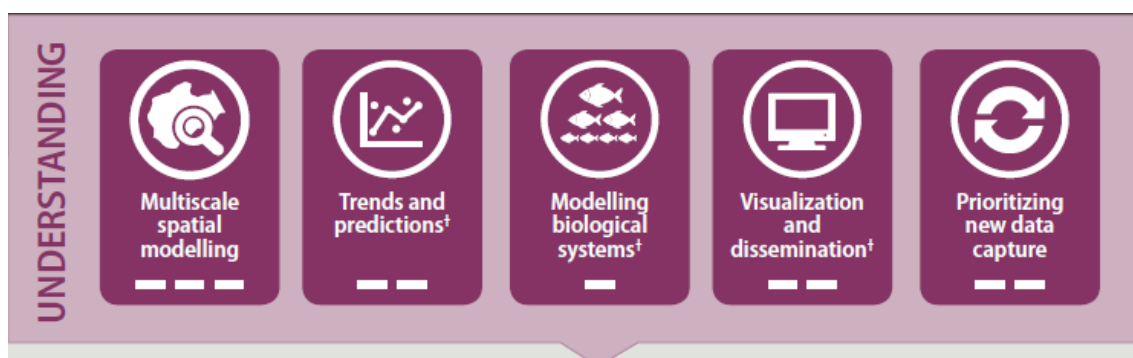


Figura 72. Cinco puntos relacionados con la capacidad de comprender. Fuente: GBIF, 2015.

Se enfoca en usar datos de biodiversidad combinados de distintas fuentes para generar información nueva, informar a los que toman las decisiones y a los que realizan las políticas y a los que ayudan a formar a la sociedad en su conjunto con la finalidad de mejorar la manera en la que se gestionan los recursos de la Tierra.

Este área persigue resolver los significativos retos existentes a la hora de generar modelos y análisis con base en todas estas fuentes de datos y recursos, de tal manera que se puedan llegar a comprender sus patrones y funciones, y con la finalidad de que puedan no solo contribuir al desarrollo de todas las demás áreas de investigación, sino también a la planificación y al desarrollo de las políticas.

Incluye, asimismo, el trabajo de analizar los datos de distribución de especies en el contexto de otras fuentes de datos geoespaciales, para aconsejar sobre posibles cambios en el tiempo y para incorporar el entendimiento de los sistemas biológicos dentro de estos modelos. También forma parte de esta área el trabajo para entregar visualizaciones informativas de esta información compleja y para identificar prioridades para futuras inversiones en movilizaciones de datos.

Los elementos que la integran son:

- *Modelización multiescala espacial:* integración de los datos recabados en las distintas disciplinas y combinación de los mismos mediante teledetección y conjuntos de datos de sistemas de información geográfica (SIG) para crear la mejor foto de la distribución geográfica de las especies.
- *Tendencias y predicciones:* integración de los datos históricos y los cambios que se producen en el tiempo con la finalidad de crear herramientas predictivas de modelización y soportar de esta manera la toma de decisiones, realizar estimaciones de biodiversidad y predecir el posible impacto derivado de los cambios de condiciones en cualquier parte de la Tierra.
- *Sistemas de modelización biológica:* construcción de modelos virtuales — desde el nivel molecular hasta los ecosistemas en su integridad— para mejorar la comprensión de los sistemas biológicos e integrar ese conocimiento en otros modelos.
- *Visualización y diseminación:* proporcionar herramientas para que la información sobre biodiversidad sea accesible y comprensible para una diversidad de audiencias, mejorando la literatura sobre biodiversidad entre el público y los encargados de realizar las políticas.
- *Priorización de la captura de nuevos datos:* el uso de los datos acumulados para identificar y priorizar nuevas oportunidades para recabar nuevos datos y proporcionar respuestas adecuadas en el tiempo a los cambios de los patrones de biodiversidad.

Aunque lo que se aborda en el GBIO no es necesariamente novedoso, aporta, sin duda, una sistematización del análisis que no solo afecta a los puntos a analizar, sino

también al nivel de desarrollo de cada uno de ellos. De esta manera, es posible centrarse en trabajar las áreas en las que se ha identificado menor desarrollo como pueden ser la modelización de sistemas biológicos (todavía muy poco desarrollada), o bien la visualización y distribución. Con carácter general, parece que el área menos desarrollada es la de comprensión, es decir, la que integra la interpretación de todos los datos recogidos con la finalidad de empezar a modelizar y a dar soluciones a las cuestiones planteadas.

El GBIO es, posiblemente, un punto de partida, una foto que sirve de arranque para aunar esfuerzos con una dirección: reducir las pérdidas de biodiversidad. Es cierto que ha habido muchas iniciativas antes, pero tal vez esta sea la que reúne toda la información hasta la fecha. Permite, de esta forma, replantear una visión a medio y largo plazo, identificando los temas sobre los que será necesario trabajar de forma sistemática, por lo que constituye un paso más. De alguna manera, el GBIO construye un marco general. En el siguiente apartado se procederá a analizar su aplicación a la realidad.

Capítulo 3. El Grupo de Observación de la Tierra en Biodiversidad

1. ANTECEDENTES

El GEO, en su plan de ejecución decenal, se estructuró en nueve áreas de beneficio para la sociedad. Una es el área de biodiversidad y, con la intención de desarrollarla, se crea GEO BON (Group on Earth Observations Biodiversity Observation Network) (GEO, 2010). Existe formalmente desde el año 2008, después de muchos años de discusión y de diseño del grupo (Scholes *et al.*, 2012).

Conectó distintas redes ya existentes, cada una de las cuales cubrió aspectos concretos de biodiversidad en diversas partes del mundo y tomó medidas para ayudar a superar las deficiencias del sistema. Se centraba en la coordinación y armonización de los sistemas existentes o en vías de crearse, la defensa y acción para sustentar los sistemas de observación y para superar las deficiencias conocidas y el comprender y dar servicio a las necesidades de los usuarios en lo que respecta a las observaciones de biodiversidad, en particular, en relación con la elaboración de políticas.

Fundamentalmente orientado a la biodiversidad,⁷³ una serie de razones propiciaron su nacimiento e hicieron posible su existencia. En primer lugar, el hecho de que la biosfera terrestre era un sistema tan complejo que no se había construido ninguna red de seguimiento que permitiese mantener un registro simultáneo de las especies individuales y de las poblaciones o un seguimiento de las tendencias en los bosques y en otros ecosistemas. Además, para presentar una foto completa de lo que estaba ocurriendo en relación con la diversidad biológica, esta red de seguimiento también necesitaría integrar cantidades ingentes de información biológica con datos y previsiones de cambio climático, polución y otras amenazas de la biodiversidad. El hecho de que no existiera una información integral en relación con los recursos biológicos mundiales seguía afectando a los esfuerzos de los responsables de las

⁷³ Disponible en: http://www.earthobservations.org/GEO_BON_a.shtml. (Consultado por última vez el 14 de marzo de 2018).

políticas y de los gestores para establecer las prioridades, elaborar las estrategias y evaluar la efectividad de sus acciones.

En segundo lugar, las nuevas tecnologías estaban mejorando de forma espectacular las colecciones y el análisis de la información relacionada con la biodiversidad. Estos sistemas cada vez más sofisticados de seguimiento que consisten en instrumentos por satélite, aire, tierra y océanos están siendo vinculados a través del GEO para conseguir crear un Sistema Global de Sistemas de Observación de la Tierra (GEOSS).

El brazo de la biodiversidad de este creciente «sistema de sistemas», llamado BON por sus siglas en inglés Biodiversity Observation Network, en español, Red de Observación de la Biodiversidad, representaría el primer paso para la consecución de una mejor comprensión de estatus y tendencias en los recursos vivos mundiales. Al poner en común los distintos instrumentos de observación autónomos y los sistemas que ahora están monitorizando las tendencias en los recursos genéticos mundiales, en las especies y en los ecosistemas, el GEO BON hizo posible crear una plataforma global que integrase los datos de biodiversidad con los datos de clima y otras variables clave. Colmó lagunas en relación con la información taxonómica y biológica y aceleró el ritmo al que se recababa y diseminaba la información.

Su objetivo era, también, verificar las exigencias de los datos de los grupos de usuarios, revisar y priorizar la investigación, facilitar la interoperabilidad entre los sistemas de observación y las bases de datos, generar de forma regular evaluaciones actualizadas de tendencias de biodiversidad, diseñar sistemas de soporte de la información que integrasen el seguimiento con modelización ecológica y previsiones y hacer que los datos y los informes estuvieran disponibles para los usuarios a través del GEOSS.

Como Red de Observación de la Biodiversidad del GEO, pretendía servir para poner en relación y, por ende, para ponderar los esfuerzos de los países, de las organizaciones internacionales y de los individuos para recabar, gestionar, compartir y analizar observaciones del estatus y de las tendencias de la biodiversidad en el mundo. Contribuirá, de este modo, a mejorar la gestión de los recursos naturales sobre el terreno y será también de utilidad a los otros ocho grupos de beneficios para la sociedad del Grupo de Observación de la Tierra: agua, salud, desastres,

clima, tiempo, energía, ecosistemas y agricultura, puesto que cada uno de ellos afecta y se ve afectado por la biodiversidad.⁷⁴

La motivación del GEO BON (GEO, 2012) radica en la necesidad crítica de que exista información sobre biodiversidad a múltiples escalas, desde lo local a lo global. Facilitando y poniendo en común los esfuerzos de los distintos países, de las organizaciones internacionales y de los individuos, el GEO BON quería contribuir a crear una recopilación, una gestión, un intercambio y un análisis de datos eficientes y eficaces en lo que respecta a las tendencias y estado de la biodiversidad a nivel mundial. De igual manera, podrá identificar los vacíos que existan en los sistemas de observación y promover mecanismos para cubrirlos.

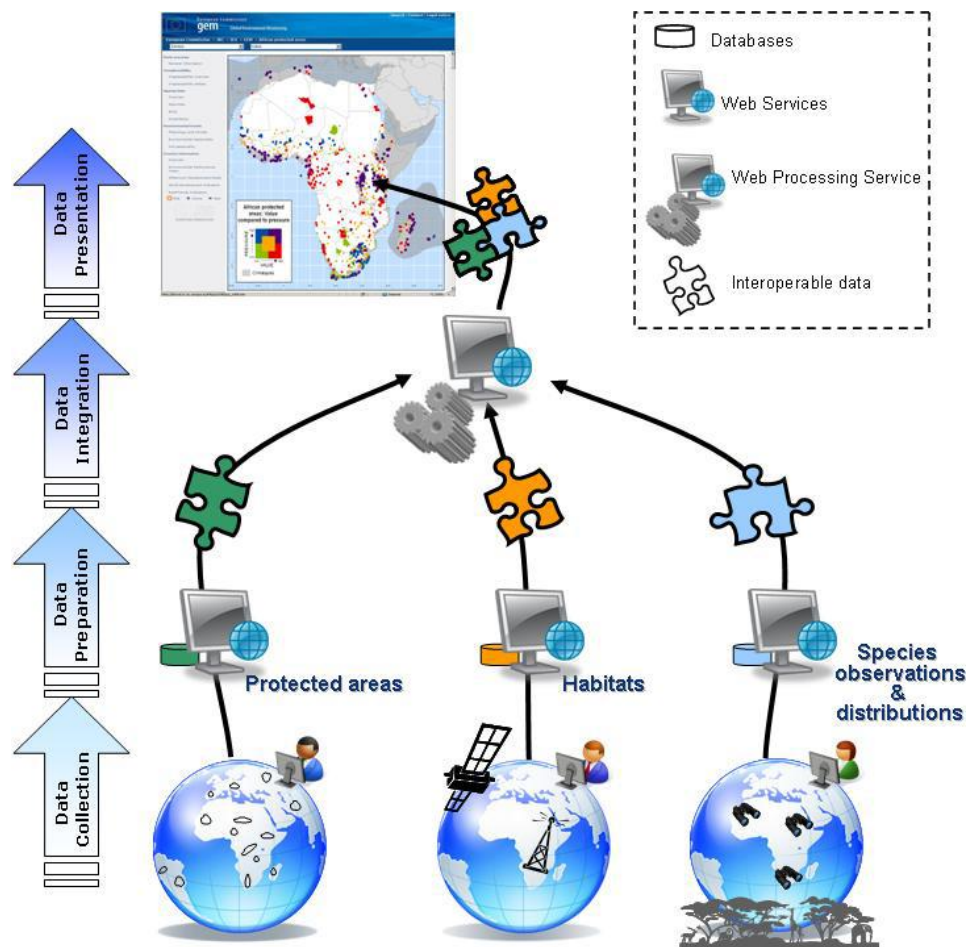


Figura 73. Partiendo tanto de las observaciones de la tierra y en remoto como de los indicadores medioambientales, es necesario recoger los datos, dejarlos listos para su uso conjunto e integrarlos para preparar los indicadores. Fuente: Dubois *et al.*, 2010.

⁷⁴ Disponible en:

http://www.earthobservations.org/documents/geo_v/07_%20GEO%20Bon%20-%20Implementation%20Overview.pdf. (Consultado por última vez el 15 de marzo de 2018).

Partiendo de la observación en el campo y de las teleobservaciones de indicadores medioambientales, es necesario recoger datos, trabajarlos para permitir que puedan usarse combinándolos e integrarlos para la preparación de indicadores (Dubois *et al.*, 2010).

La definición de biodiversidad que recoge el GEO BON es consistente con la del CDB e incluye variación en la composición, estructura y función de los ecosistemas, especies y niveles genéticos de organizaciones biológicas.

El alcance del GEO BON incluye la biodiversidad terrestre, la de agua dulce, la de las aguas costeras y la de los medios marinos. Además de recoger series de tiempo de observaciones relacionadas con la presencia, abundancia y condiciones de los elementos de biodiversidad en todos estos ámbitos, y en todos esos medios, en última instancia también recoge información sobre interacciones entre organismos, el uso que de los mismos hacen las personas y los datos que soportan el entorno abiótico, la situación taxonómica, los motores del cambio de la biodiversidad y las medidas adoptadas para proteger la biodiversidad. También da soporte a evaluaciones más detalladas realizadas por parte de organismos de evaluación de biodiversidad y ecosistemas.

La parte de ecosistemas del GEO BON proporciona información global en relación con ecosistemas terrestres, marinos y de agua dulce, centrándose en su distribución, extensión y condiciones y en la manera en que dichos parámetros están cambiando en el tiempo. El motor de estos cambios y las consecuencias de los mismos son especialmente interesantes, como también lo son los vínculos entre los distintos ecosistemas.

La parte de especies se centra en varios aspectos críticos de las especies. El primero de ellos es de qué manera la distribución de las especies y la abundancia de las mismas está cambiando. Se va a realizar una evaluación usando un cuidadoso esquema de ejemplo global, utilizando especies representativas.

El segundo aspecto crítico es facilitar la creación de mapas de distribución para un gran número de especies y el acceso a ellos. Esto va a servir para incrementar el trabajo ya empezado por parte de un gran número de organizaciones. La modelización de especies y de comunidades se integra con series de datos cronológicas con la finalidad de contribuir a evaluar las consecuencias de los cambios observados en términos de impactos en distribuciones biológicas.

La diversidad genética se monitoriza con una combinación de formas de teleobservación y de observación *in situ* con trabajo en el ámbito de los ecosistemas y de las especies. Se usan varias estrategias de observación. La primera de ellas es la observación directa de compuestos genéticos concretos relativos a ciertas especies concretas. La segunda es la observación de otros componentes como puede ser la

extensión de la gama de conjuntos de especies representativos que, por lo tanto, quedan integrados en modelos de los que se infiere la diversidad genética. Y la tercera está constituida por las observaciones de condiciones de los ecosistemas — por ejemplo: inferidos mediante el uso de tele observación— integrados en los modelos de variación espaciales y genéticos sobre los que se basan las inferencias a gran escala relacionadas con los cambios a nivel genético, difíciles de obtener directamente.

Además del trabajo en ecosistemas, especies y genes, el GEO BON también se centra en muchas otras áreas, por ejemplo: integrar observaciones de ecosistemas para ayudar a evaluar el estado de los bienes, servicios y funcionamiento de los ecosistemas. La modelización es una de las primeras herramientas para contribuir a integrar las observaciones en el tiempo y en el espacio, entre las taxonomías y para realizar previsiones sobre los futuros cambios. La capacidad de construir esfuerzo fortalece la capacitación, principalmente en los países en vías de desarrollo, de recoger y utilizar observaciones de la Tierra y contribuir así al GEO BON.

Proporciona a los usuarios información que interpreta y documenta los cambios que se están produciendo en biodiversidad. Esta información sirve de base para futuras evaluaciones del IPBES. Estos sistemas están diseñados con esa finalidad.

Los principales usuarios del GEO BON son los países —en particular, en lo que respecta a sus obligaciones al amparo de las convenciones relacionadas con la biodiversidad y a los recursos naturales—, las agencias de conservación de la biodiversidad y las organizaciones internacionales

Existe un acuerdo sobre la necesidad de un sistema de observación de la biodiversidad general y con integración en distintos niveles, desde el ámbito regional hasta el ámbito global a los efectos de mejorar la biodiversidad y el bienestar del hombre. Este sistema debe contribuir a establecer comparaciones del estatus de la biodiversidad en distintos lugares y monitorizar los cambios en biodiversidad en un momento dado en el tiempo.

Los principales agentes dentro de esta empresa incluyen la biodiversidad nacional y comunidades de conservación de la naturaleza, agencias de conservación y gestión de la naturaleza, departamentos de Gobiernos nacionales responsables de las obligaciones de biodiversidad derivadas de la firma de tratados internacionales, etc.

El GEO BON no empezó de cero, pues las observaciones relacionadas con la biodiversidad se encuentran entre las observaciones registradas hace más tiempo y son, sin duda, más numerosas. Existían muchas colecciones de plantas y animales en museos y herbarios alrededor del globo, cientos de millones de observaciones de campo por parte de profesionales y legos en la materia y terabits de imágenes recogidas en remoto y mapas de los cambios en la Tierra. Por lo general, sin

embargo, las observaciones eran locales, descoordinadas, observaciones *in situ* y muy poco armonizadas a nivel global. Además, los sistemas de observación terrestres, marinos y de agua dulce no estaban interconectados y existían muchas diferencias en los enfoques de la observación y de las taxonomías. Para complicar todavía más todo esto, las observaciones registradas en remoto, por lo general, se usaban para describir la utilización del suelo y de los ecosistemas.

Han surgido toda una serie de organizaciones nacionales e internacionales con la finalidad de gestionar todos estos datos y llevarlos a la era de la interconexión, puede verse, por ejemplo, el caso del GBIF.

2. ESTRUCTURA

El GEO BON está integrado por agencias gubernamentales, intergubernamentales y organizaciones internacionales. De acuerdo con un plan de implementación actualizado, estos socios trabajan conjuntamente de forma voluntaria para coordinar y conectar sus sistemas de observación y desarrollar servicios de información y decisiones de soporte de producto.

Dirigido por un *Steering Committee*, reporta con regularidad al Plenario del GEO y a sus cumbres ministeriales. Todos los sistemas, productos y servicios a los que contribuya el GEO BON o que inicie este quedan totalmente sujetos a la autoridad y son propiedad de las agencias que los producen y gestionan.

Para alcanzar su visión está dando los siguientes pasos:

- Identificar a los proveedores de sistemas de observación, de datos y de bases de datos, de servicios de información y de otros recursos e invitarlos a contribuir dentro del marco del GEO BON.
- Construir una red de personas y de organizaciones dispuestas a colaborar y compartir ideas e información.
- Partiendo de estándares técnicos acordados para metadatos e interoperabilidad, trabajar hacia la integración de varios tipos de datos de biodiversidad con otros datos relevantes accesibles a través del GEOSS.
- Identificar lagunas de cobertura, reunir a los socios para gestionarlos y trabajar en reforzar y hacer sostenibles los sistemas existentes de seguimiento.
- Transformar los datos y la información en soporte de decisiones dirigidas al usuario y a los servicios finales.

- Promover un acceso total y abierto a los datos de biodiversidad no sensibles, de conformidad con lo previsto en los principios de distribución de los datos del GEOSS.
- Disseminar los conjuntos de datos, las herramientas de soporte y elaborar modelos a través de portales web fáciles de uso.
- Desarrollar y contribuir a los programas para construir la capacidad de los individuos y de las instituciones, en particular, en los países en vías de desarrollo.

Con base en estos puntos, ya ha empezado a coordinar la recopilación de datos y la entrega de información. Durante los próximos años, proporcionará un número creciente de información relevante de productos y servicios. El GEO BON continuará comprometiendo a las comunidades científicas en la gestión de recursos y de políticas para asegurar el crecimiento de la disponibilidad de información de biodiversidad de alta calidad.

3. OBJETIVO

El objetivo de GEO BON queda recogido en GEO-V (2008). En un primer momento, el reto del GEO BON (Scholes *et al.*, 2008) fue construir la red, localizar y contratar proveedores de datos de biodiversidad. La biodiversidad tiene mucha complejidad, y no solo la propia de la disciplina, sino también la resultante del hecho de que la observan un grupo variado de organizaciones, instituciones e individuos con muy distintas agendas, intereses, presupuestos, limitaciones e intereses.

En lo que respecta a algunos datos, como los de especies, existen ya redes fuertes y estándares que dan lugar a la existencia de plataformas robustas. Por ello, el GEO BON ha tenido que dedicar parte de sus recursos a mejorar la interoperabilidad con estándares, herramientas, etc. Además, a resultas de la importancia de poder compartir los datos, procedió a diseñar tanto las observaciones *in situ* como las remotas en tres niveles de interés para los investigadores de biodiversidad y para los que redactan las políticas: genética, especies y ecosistemas. Por lo general, cada tipo y nivel de organización biológica tiene sus propios datos y metadatos estándares que proceden de distintos proveedores de redes.

El GEO BON construye tomando como base las experiencias de varios socios y suministradores de datos, como pueden ser el GBIF y el GEOSS ponderando mecanismos que estén en marcha y añadiéndoles aquellos elementos que permiten la integración de datos biológicos *in situ* y remotos.

Otra de sus funciones importantes es facilitar las interacciones entre participantes para generar sinergias. El GEO BON puede verse como un sistema dinámico de

sistemas para observar, almacenar y compartir datos, pero dicho sistema deben construirlo personas, que trabajen en organizaciones, instituciones, Gobiernos o incluso para sí mismas, para lograr sus objetivos personales. Como las personas coinciden en redes científicas, es importante coordinar la estructura que funcione para que se integren las acciones separadas para lograr los objetivos que decidan en común.

Además, el GEO BON tiene que tener la capacidad de construir una estrategia que permita el crecimiento, el mantenimiento y el equilibrio dentro de su red. Esto estará basado en los principios que articulan el plan de implementación del GEOSS.

La construcción del GEO BON debía desarrollarse de forma consultiva entre la red de participantes e incorporar acciones para fortalecer la capacidad de todos los países miembros del GEO, en particular, de los países en vías de desarrollo, para que pudieran utilizar los datos de observación de la Tierra y los productos y contribuir a las observaciones y sistemas del GEO BON. También debía facilitar la identificación de soluciones que potenciasen la sostenibilidad de modo que mejorase la contribución de la ciencia a la sociedad.

El GEO BON es una red de organizaciones de la que forman parte:

- Los convenios relacionados con la biodiversidad (CITES, CDB, etc.).
- Las organizaciones de Naciones Unidas.
- Las organizaciones nacionales e internacionales de conservación.
- Las agencias y Gobiernos nacionales.
- Las instituciones académicas y de investigación.
- El sector privado.
- Las fuentes de información genética.
- DIVERSITAS (el programa internacional de ciencia de la biodiversidad de la UNESCO).
- El GBIF.
- La red ILTER.
- Las agencias espaciales.
- El Centro Mundial de Vigilancia de la Conservación.

Por consiguiente, está integrado, principalmente, por los sistemas de contribución con responsabilidad primaria para observaciones de biodiversidad en el ámbito global, nacional e infranacional. Una de sus actividades principales y más

importantes será comprometerse con varios titulares de datos de biodiversidad, aportando sobre el trabajo del GBIF y, de otros, con la finalidad de ofrecer una multiplicidad de herramientas analíticas y de modelos para la interpretación y uso de las observaciones para la biodiversidad existentes. Será fundamental realizar una consulta iterativa y un proceso de construcción y será, asimismo, imperativo partir de la experiencia de otros esfuerzos globales.

El GEO BON se beneficia y contribuye, al mismo tiempo, al GEOSS basándose en los recursos de los potenciales participantes y socios del GEO BON. Los potenciales participantes son instituciones miembros del GEO, organizaciones no gubernamentales, suministradores de datos y agregadores, desarrolladores de herramientas y operadores, así como otro tipo de profesionales que tengan algún tipo de interés en las observaciones de biodiversidad y sus aplicaciones a la conservación y a la gestión mejorada de recursos naturales.

Uno de los primeros pasos para su implementación fue la temprana integración de algunos productos útiles. Estos productos fueron ejemplos de los tipos de análisis que el GEO BON generaría e incluiría como pueden ser los mapas alineados de extensiones de ecosistemas, y las distribuciones de datos de especies (GEO BON, s. f.). Las personas que tomaban decisiones tenían que poder confiar en la ciencia a la hora de diseñar políticas efectivas y programas para conservar y gestionar los cada vez más limitados recursos biológicos.

Se han ido implantando una multiplicidad de sistemas de observación de biodiversidad y de bases de datos a lo largo de los años, pero estos sistemas han tendido a estar dispersos, desconectados y descoordinados. Con el paso de los años, han aparecido barreras técnicas e institucionales que pueden dificultar el acceso a las observaciones y datos recogidos. Existen también lagunas de cobertura y de datos de integración biológica con datos de otras disciplinas, como puede ser el clima, el tiempo y geología, por lo que sigue siendo difícil.

El GEO BON está empezando a coordinar la provisión de datos e información de biodiversidad, sostenible, integrada, intersectorial y accesible. Habida cuenta de que el seguimiento y la gestión de los datos de biodiversidad tienen un coste alto, la comunidad científica tiene un gran incentivo para trabajar asociándose, explotar las sinergias y compartir los datos. Facilita esta cooperación, proporcionando un marco global y científicamente sólido para observaciones a largo plazo y para detectar cambios en la biodiversidad, pero no existe un mecanismo de opinión científica consensuada.

4. EL OBSERVATORIO EUROPEO DE BIODIVERSIDAD

Aún hoy existen lagunas clave en lo que respecta al estatus y a las tendencias relacionadas con la biodiversidad y los servicios asociados a los ecosistemas, fundamentalmente como resultado de las barreras que hacen que los datos que existen no puedan encontrarse, no sean accesibles y digeribles (es decir, interoperables) (Wetzel *et al.*, 2015). Los datos, todavía muchas veces, no son fáciles de localizar por estar cargados en un repositorio públicamente conocido (portal), o bien por no estar lo bastante documentados o estructurados o, incluso, por tener metadatos. Además, con frecuencia, los datos existentes no son accesibles, lo que hace que no estén disponibles para usarse, por ejemplo, en los casos en los que existan restricciones de uso (licencias) o confidencialidad. Estas cuestiones están contempladas con mayor profundidad en el punto del presente trabajo de investigación relacionado con la interoperabilidad. Por último, los datos no son digeribles, por ejemplo, porque no siguen estándares acordados y ello hace difícil integrarlos o combinarlos con otros datos similares (Wetzel *et al.*, 2015).

Parar la pérdida de biodiversidad requiere de la implementación de una multiplicidad de iniciativas. El Observatorio Europeo de Biodiversidad (en adelante, EU BON) formó parte de las iniciativas adoptadas para la consecución de este importante reto (Hoffmann *et al.*, 2014).

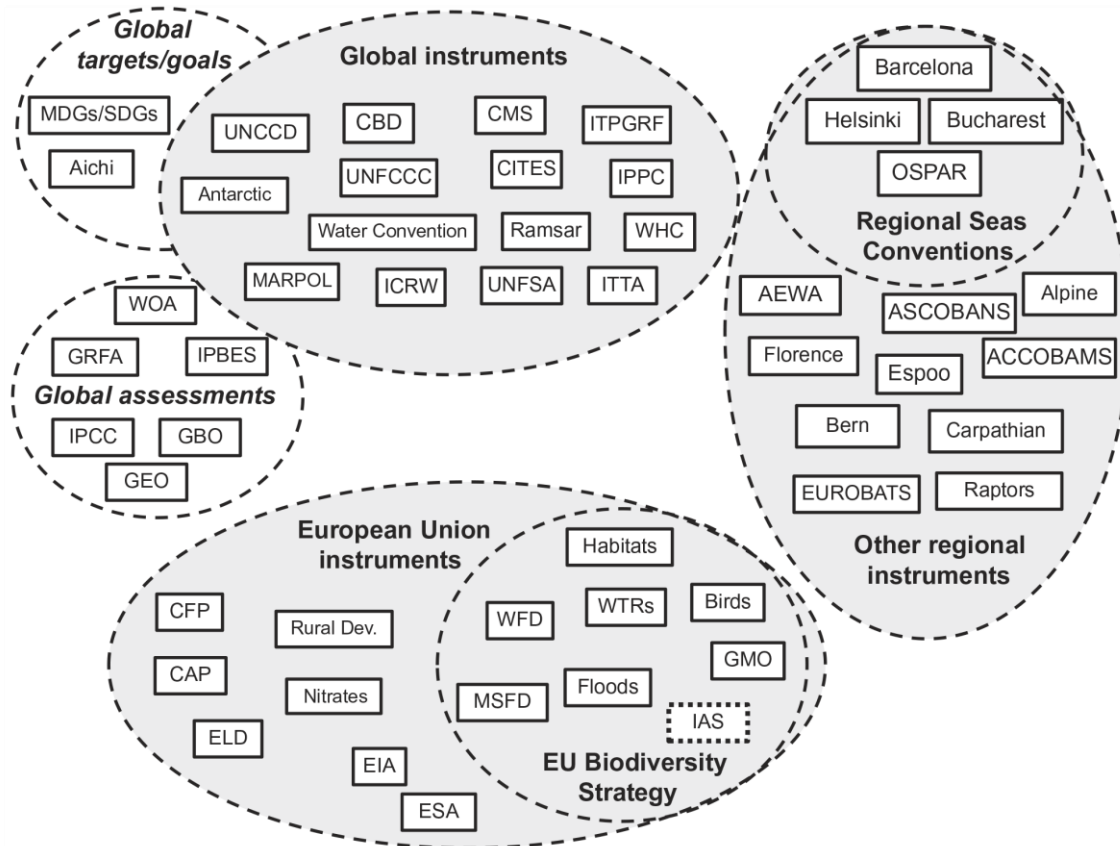


Figura 74.

Muestra de la complejidad existente en Europa. Fuente: Wetzel *et al.*, 2015.

Este proyecto se enfocó principalmente en el continente europeo, aunque contribuía al mismo tiempo a una iniciativa mucho más global, el GEO BON, que, a su vez, forma parte del GEOSS.

La idea es que el EU BON pudiera fundarse en e-infraestructuras ya existentes como el GBIF y LifeWatch, así como los centros de datos de biodiversidad nacionales existentes en Europa, integrando los datos de biodiversidad relevantes tanto de las observaciones *in situ* como de las teleobservaciones, relacionadas con los hábitats terrestres, marinos y de agua dulce. Una de las claves del EU BON era la entrega de datos relevantes, totalmente integrados a una multiplicidad y variedad de *stakeholders* y usuarios finales tanto de ámbitos locales como globales.

Desarrollar y aplicar nuevos estándares y protocolos era clave para construir una mayor interoperabilidad entre las distintas capas y sistemas, proporcionar un acceso a herramientas y servicios analíticos mejorados, y un mejor y más armonizado registro y sistemas de control de los esfuerzos de la ciencia ciudadana⁷⁵ con la finalidad de lograr programas de investigación a largo plazo que permitieran incorporar a futuro los datos recogidos. Sirve como soporte para las interfaces de biodiversidad científico políticos, facilitando la toma de decisiones, permitiendo una toma de decisiones en el ámbito de la biodiversidad fundadas y contribuyendo a conservar la biodiversidad para el bienestar humano a distintos niveles. El proyecto quiso reforzar las capacidades europeas y las e-infraestructuras para la gestión de la información medioambiental y el desarrollo sostenible.

⁷⁵ Se entiende por ciencia ciudadana a la investigación científica llevada a cabo por una suma de colaboradores, en su totalidad o en parte por científicos y profesionales junto a gente común. Formalmente, la ciencia ciudadana ha sido definida como «la recopilación y análisis sistemático de datos, el desarrollo de la tecnología, las pruebas de los fenómenos naturales, y la difusión de estas actividades por los investigadores sobre una base principalmente vocacional».

El documento *Green Paper on Citizen Science: Citizen Science for Europe* describe la ciencia ciudadana como «el compromiso del público general en actividades de investigación científica; cuando los ciudadanos contribuyen activamente a la ciencia con su esfuerzo intelectual o dando soporte al conocimiento con sus herramientas o recursos. Los participantes proveen datos experimentales o equipos a los investigadores. Los voluntarios, a la vez que aportan valor a la investigación, adquieren nuevos conocimientos o habilidades, y un mejor conocimiento del método científico de una manera atractiva. Como resultado de este escenario abierto, colaborativo y transversal, las interacciones entre ciencia-sociedad-políticas investigadoras mejoran, conduciendo a una investigación más democrática, basada en la toma de decisiones basada en evidencias informadas surgidas del método científico, total o parcialmente, por parte de científicos amateur o no profesionales». Ciencia Ciudadana (s. f.). En Wikipedia. Recuperado de https://es.wikipedia.org/wiki/Ciencia_ciudadana. (Consultado por última vez el 9 de diciembre de 2016).

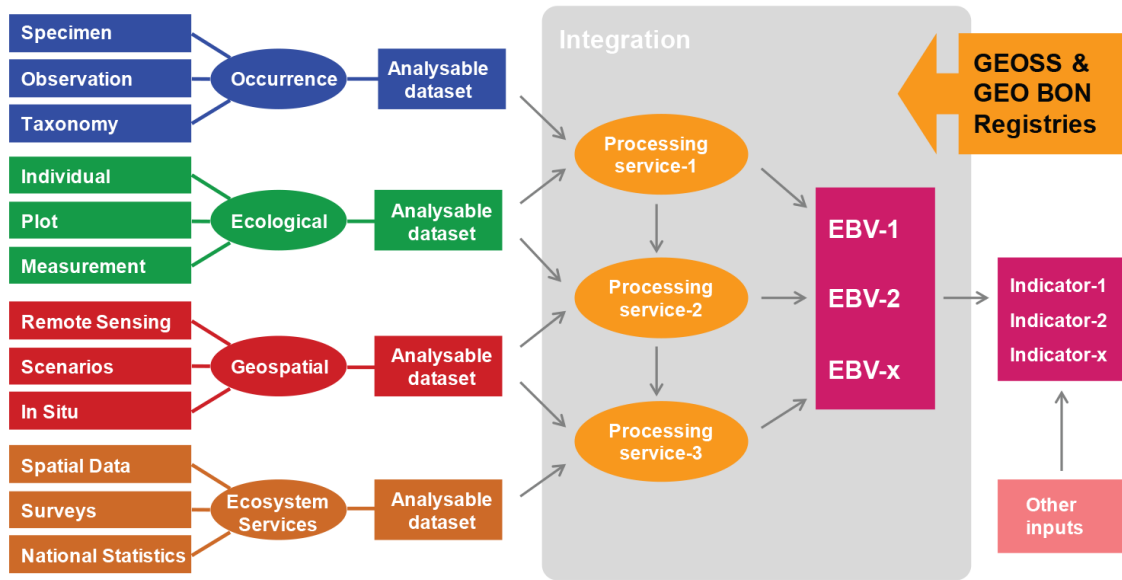


Figura 75. El EU BON y el desarrollo de la visión del GEO BON. Fuente: Hoffmann *et al.*, 2014.

El objetivo previsto del EU BON fue, como señalan Hoffmann *et al.* (2014), usar su potencial para modificar la interrelación entre ciudadanos, ciencia y normativa para la biodiversidad. Los que tomaban las decisiones a distintos niveles debían ser capaces de utilizar la información de biodiversidad adaptada a sus necesidades concretas y las bases de datos dispares y desconectadas, poder integrarse con la finalidad de permitir hacer un seguimiento y una evaluación de medidas en ámbitos espaciales y temporales distintos. Esto requiere ímprobos esfuerzos no solo respecto de las armonizaciones técnicas de las bases de datos, de los modelos y de las herramientas de visualización, sino también en el diálogo de las redes sociales asociadas, cubriendo una variedad de organizaciones científicas y de ciencia ciudadana.

Tal vez, el EU BON fuese un ejemplo de una interfaz científico-política de menor envergadura, es decir, no global (Soberón y Sarukhán, 2010). Sin embargo, parece que está adoptando una dinámica propia, mientras que el GEO BON no es otra cosa que la aplicación de la informática del GEO a la biodiversidad.

Capítulo 4. La aplicación en una escala inferior a la global, dos modelos continentales de gestión de la bioinformática: Europa y Estados Unidos. Otros ejemplos de modelos continentales

1. EL MODELO EUROPEO: LIFEWATCH ERIC (E-SCIENCE AND TECHNOLOGY INFRASTRUCTURE FOR BIODIVERSITY DATA AND OBSERVATORIES)

De la aplicación de la teoría de los ERIC a las ciencias, a la biodiversidad, nació LifeWatch como parte del *roadmap* de los ESFRI; aparece en la parte dedicada a las ciencias medioambientales, por lo que queda enmarcada entre las infraestructuras de investigación de EUDAT (European Data Infrastructure) (Lossau, 2012). Una de las principales características de EUDAT es lo multidisciplinar de los datos que gestiona.

La fase preparatoria empezó en 2011 y el año previsto de operación era 2014. Sin embargo, una serie de lagunas fueron dificultando el arranque de LifeWatch (LifeWatch, 2015):

1. La diferencia existente entre la práctica actual y la visión a futuro. En la actualidad, prevalece una visión reduccionista de la investigación relacionada con la biodiversidad. Se hace necesario realizar un cambio de paradigma que permita descubrir los patrones y definir los procesos subyacentes. LifeWatch debe servir para favorecer este nuevo enfoque permitiendo analizar enormes conjuntos de datos, con un amplio rango de escalas.
2. La heterogeneidad de los requerimientos de la comunidad, las fuentes de datos y herramientas y la manera en la que lidiar con la complejidad de la naturaleza interconectada de las ideas de la biodiversidad y de los ecosistemas, de los resultados y de los repositorios.

3. La fragmentación de la comunidad derivada del gran número y alcance de los actores que intervienen (curadores, investigadores, gestores, compañías, desarrolladores de aplicaciones), hábitats (marino, terrestre, de agua dulce, etc.) y de escalas de interés (organización biológica, función, espacio, tiempo, etc.). LifeWatch debe contribuir a una fuerte reducción de dicha fragmentación y funcionar como soporte de la investigación multidisciplinar proporcionando un conjunto de e-servicios común y promoviendo la adopción de estándares.
4. La fragmentación de los datos. En los últimos años, un gran número de proyectos europeos y nacionales han generado un alto número de datos, información y conocimiento en los ámbitos de la biodiversidad y de los ecosistemas. Sin embargo, esta información está fragmentada. LifeWatch debe contribuir a la mejora de la accesibilidad a los datos dispersos y fragmentados y proporcionar los elementos de análisis necesarios para descubrir nuevos patrones dentro de los sistemas complejos. Por consiguiente, en su condición de infraestructura de investigación estable y sostenible, se espera que fomente la investigación en el ámbito de las ciencias medioambientales.
5. La necesidad de que exista interoperabilidad dentro del amplio espectro de datos, información y repositorios de conocimiento y también teniendo en cuenta la multiplicidad de ideas y resultados proporcionados por las aplicaciones de modelización. Como tal, LifeWatch debe poner especial atención en las recomendaciones de la Research Data Alliance (RDA).⁷⁶
6. El desarrollo de nuevas aplicaciones de modelización. Hoy en día, la integración de las aplicaciones existentes muestra que los modelos en los que están basadas no están lo bastante bien interconectados, no están suficientemente alimentados ni por la cantidad ni por la calidad de datos, de manera que no proporcionan suficiente síntesis y comprensión. Por consiguiente, tienen que desarrollarse nuevas aplicaciones de modelización con la finalidad de lidiar con la complejidad asociada a los datos de biodiversidad (ecosistemas, especies y genes), a las funciones de biodiversidad, a los instrumentos necesarios para llevar a cabo este estudio (colecciones, observatorios, herramientas de modelización y de análisis) y sus conexiones.
7. La escala de implementación de e-infraestructuras paneuropeas. La conexión con los usuarios y los generadores de datos a través de enlaces

⁷⁶ Es una organización fundada por la Comisión Europea, la National Science Foundation de Estados Unidos y el Gobierno australiano. Para conocerla con más detalle, puede consultarse <https://www.rd-alliance.org/about-rda>.

externos con otras infraestructuras de investigación como puede ser ELIXIR, ICOS (por las siglas en inglés de Integrated Carbon Observation System, Sistema Integrado de Observación del Carbono), EMSO (European Multidisciplinary Seafloor and Water Column Observatory), European contribution to ARGO Programme-EURO ARGO, JERICO-Next, entre otras. Y también, con otras grandes redes de excelencia que, a cambio, también pueden agrupar otras iniciativas nacionales relevantes, en los ámbitos de investigación oceánica, marina, terrestre, etc.

8. El ritmo acelerado de innovación en nuevas tecnologías con la finalidad de evitar que LifeWatch pueda quedar obsoleto; será necesario hacer un seguimiento de la tecnología.
9. El encaje con la industria general y la formación e investigación en el sector.

Después de una serie de retrasos en la constitución de LifeWatch, el proyecto se ha vuelto a poner en marcha. La idea es que LifeWatch proporcione un único punto de acceso común para los (ciudadanos) científicos y los que elaboran las políticas con la finalidad de que puedan encontrar, analizar y buscar datos nuevos y disponibles.

Una arquitectura ágil basada en la tecnología de la información procesamiento distribuido abierto (OGC), estándares abiertos distribuidos (ODP) y los estándares de cumplimiento que establece la directiva INSPIRE proporciona (Frenzel *et al.*, 2011): recursos como pueden ser los repositorios de datos, y las redes con capacidad computacional y de observación (sensores) conectadas para que las e-infraestructuras permitan compartir esos recursos en Europa, facilitando la creación de *kits* de herramientas, flujos de datos y conjuntos de datos. De ellos resultará una infraestructura orientada al usuario con e-servicios o e-laboratorios o incluso laboratorios virtuales, donde los usuarios puedan colaborar, comunicarse y experimentar entre los distintos países y las disciplinas de investigación de diversas maneras.

LifeWatch construye partiendo de las redes existentes de datos. Se aprueba la Decisión de Ejecución (UE) 2017/499 de la Comisión de 17 de marzo de 2017 relativa a la creación de la Infraestructura Virtual Europea de Ciencia y Tecnología para la Investigación sobre la Biodiversidad y los Ecosistemas —Consorcio de Infraestructuras de Investigación Europeas (LifeWatch ERIC)— (Comisión Europea, 2017). Con esta decisión, por fin, LifeWatch se pone en marcha.

Involucra a ocho países europeos: Italia, Grecia, Bélgica, Portugal, Eslovenia, España, Rumanía y Holanda. Los ERIC para su constitución necesitan un mínimo de tres Estados (no necesariamente los tres necesitan ser europeos, pero la mayoría deberán ser europeos). La idea de base fue que la coordinación de LifeWatch la llevara a cabo la Universidad de Ámsterdam. Al final, la coordinación se ejecutará

desde Andalucía. Esta e-infraestructura reúne ocho redes europeas de biodiversidad y diecinueve Gobiernos nacionales, a través de comités, otros usuarios y plataformas (Hernández Ernst *et al.*, 2009).

El presupuesto de construcción de la infraestructura LifeWatch está estimado en doscientos veinte millones de euros. España, más concretamente Sevilla, será la sede administrativa y legal, además de albergar las instalaciones centrales relacionadas con las tecnologías de la información y comunicación, basándose en las instalaciones existentes de la Reserva Biológica de Doñana (CSIC). Allí se encuentra fijada la sede estatutaria del ERIC, donde este se aloja. Se encuentra allí también la coordinación y gestión de las operaciones así como las infraestructuras tecnológicas. Holanda acogerá el Centro TIC de Investigación e Innovación de LifeWatch, los laboratorios virtuales, es decir, toda la gestión del *middleware* y servicios a usuarios, mientras que Italia alojará su Centro de Servicios (González Aranda, 23 de noviembre de 2011).

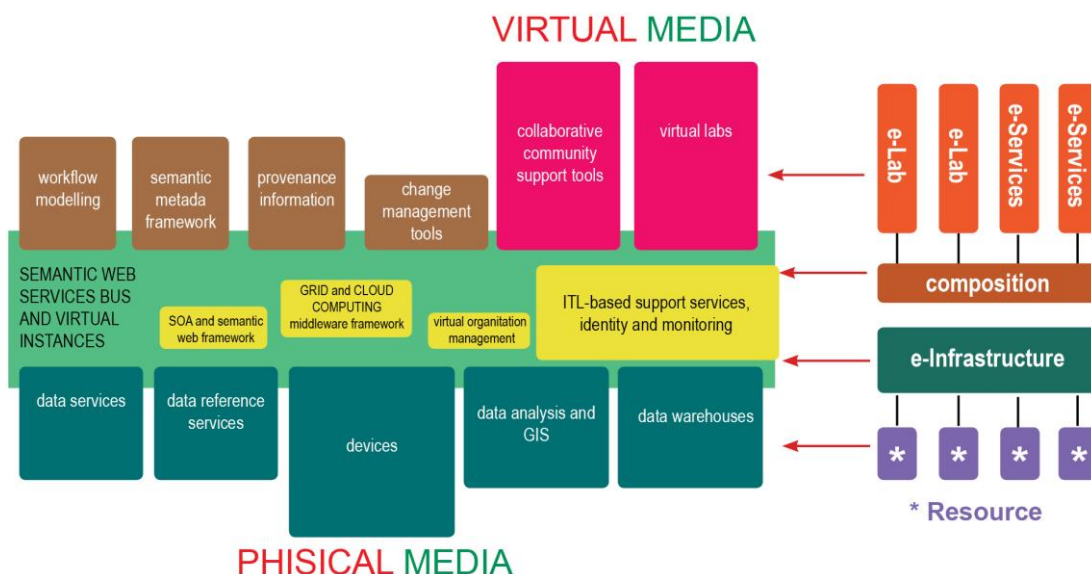


Figura 76. Actividades principales realizadas en las distintas capas de la e-infraestructura de LifeWatch. Fuente: LifeWatch, 2015.

Es e-infraestructura distribuida que permitirá la conectividad de las bases de datos de biodiversidad y ambientales relacionadas con ella. Tiene vocación de convertirse en una referencia mundial para la protección, gestión y uso sostenible de la biodiversidad. Quiere construir la infraestructura de e-Ciencia y Tecnología de Investigación para Datos y Observatorios relacionados con la Biodiversidad. Reúne:

1. Una red de observatorios marinos, terrestres y de aguas dulces.
2. Acceso único y abierto a una gran cantidad de información distribuida en bases de datos interconectadas y sitios de monitorización.

3. Instalaciones de computación en laboratorios virtuales con herramientas analíticas y de modelización.
4. Apoyo y formación a usuarios identificados y un programa de servicios públicos.

La Primera Asamblea General se celebró en Sevilla los días 8 y 9 de mayo de 2017. El consorcio español está formado por el Ministerio de Economía y Competitividad, la Junta de Andalucía —a través de la Consejería de Economía y Conocimiento y la Consejería de Medio Ambiente— y la Confederación Hidrográfica del Guadalquivir. La iniciativa tendrá la sede central en Sevilla y, como tal, ha iniciado su andadura (Europa Press, 2017).

La Junta participó en la Primera Asamblea del proyecto europeo LifeWatch. Está basada en la creación de una infraestructura electrónica destinada a reforzar la capacidad científica, tecnológica y de innovación en el ámbito de la biodiversidad. Su objetivo es establecer una gran red de datos, información y conocimiento basada en una potente e-infraestructura de comunicaciones, supercomputación y *cloud* distribuida (*big data*). Esta iniciativa brindará a la comunidad científica, a los gestores ambientales y a los ciudadanos el acceso a grandes volúmenes de datos relacionados con la biodiversidad (bióticos: fauna, flora; abióticos: medios atmosférico, terrestre, oceánico y fluvial o aguas dulces) y permitirá trabajar sobre ellos con herramientas de análisis, disponibles de manera virtual (Europa Press, 2017).

La primera reunión operativa ha tenido lugar el 19 de enero de 2018 y ha servido para aprobar el plan estratégico.

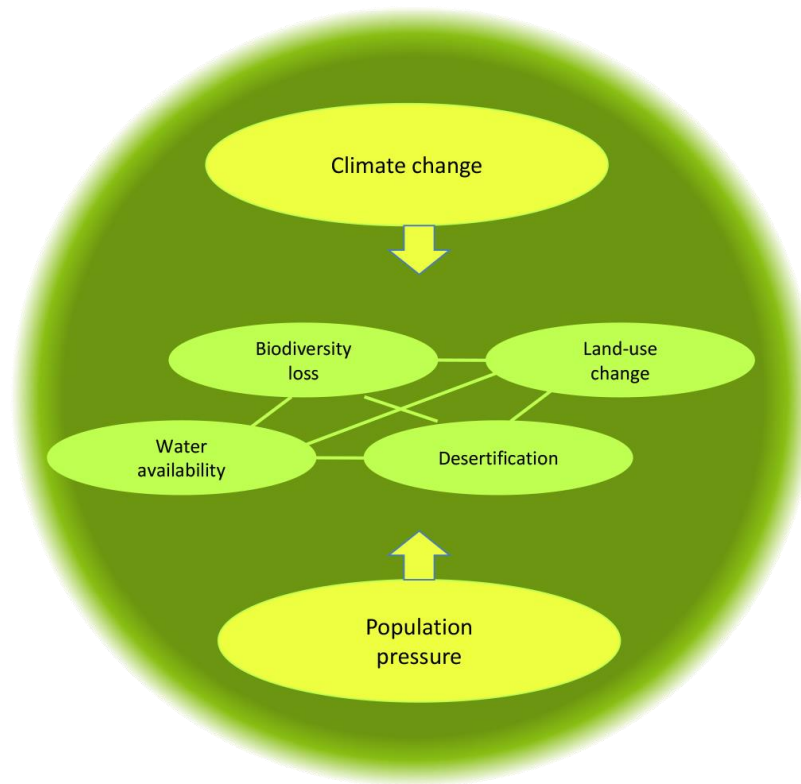


Figura 77. Impacto antropológico. Fuente: LifeWatch, 2015.

Dentro de este marco y de conformidad con el *roadmap* del Foro de Estrategia Europa de Infraestructuras de Investigación, LifeWatch, la infraestructura europea de e-ciencia para la biodiversidad y la investigación de ecosistemas quiere dar soporte a los investigadores dando acceso a los laboratorios virtuales, permitiendo analizar y modelizar los datos de biodiversidad existentes o nuevos con herramientas de bioinformática avanzadas. Los investigadores pueden colaborar en estos laboratorios virtuales para desarrollar conocimiento de biodiversidad en un entorno multidisciplinar.

Como infraestructura de investigación, la arquitectura de LifeWatch debe servir para proporcionar a la comunidad científica capacidades avanzadas y rápidas que permitan dar respuesta a cuestiones complejas dentro del ámbito del sistema (Basset y Loss, 2012). Esto es posible teniendo en cuenta que existen varios elementos clave ya en marcha como pueden ser las relacionadas con la diversidad genómica (por ejemplo: GenBank, www.ncbi.nlm.nih.gov/genbank), la diversidad de especies (por ejemplo: Atlas of Living Australia, www.ala.org.au; y World Register of Marine Species (WORMS), www.marinespecies.org), diversidad de ecosistemas y hábitat (por ejemplo: European Nature Information System (EUNIS), <http://eunis.eea.europa.eu>), etc.

Quiere promover la integración de datos de medioambiente y biodiversidad de otras e-infraestructuras (el GBIF, el GEOSS, etc.). Es una de las claves con las que

Europa va a contribuir al GEO BON (Hardisty *et al.*, 2013). El desarrollo de este tipo de infraestructuras de biodiversidad es incluso una recomendación dentro de la Unión Europea para la consecución de los fines de preservación de la biodiversidad entre otros (De Meester *et al.*, 2011); este tipo de e-infraestructuras son esenciales para contribuir a frenar la pérdida de biodiversidad.

Es, además, un marco para acceder a entornos virtuales integrados en los que se pueda compartir la aplicación de trabajos de análisis integrados, etc. Permitir a los científicos trabajar en entornos virtuales compartidos no solo contribuye a mejorar la colaboración y difundir las buenas prácticas, sino que, además, facilita la transferencia del conocimiento de los científicos a las personas que elaboran las normas (*policy makers*).

Permite, asimismo, que estas personas puedan acceder a mucho más nivel de detalle en ese mismo entorno virtual. Esta infraestructura, además, quiere impulsar la relación entre científicos y ciudadanos con la finalidad de ampliar la difusión de las prácticas de los científicos. La implicación de los ciudadanos es clave: el ciudadano tiene que dejar de ser un mero espectador y pasar a implicarse (Santamaría, 2017).

En LifeWatch los responsables políticos pueden trabajar en estos problemas directamente o con la ayuda de los investigadores locales en plataformas de investigación activa en una interfaz científico-política. La política ambiental puede pasar de la incertidumbre a una mayor confianza y seguridad.

LifeWatch quiere crear nuevas herramientas que permitan el análisis y la modelización de bases de datos grandes y heterogéneas. Entre otros aspectos, se hace posible la identificación de relaciones para esclarecer los procesos subyacentes relacionados con el análisis de la observación de los fenómenos derivados del cambio global, y también abrir nuevas oportunidades para la experimentación orientada a un objetivo.

La creación de LifeWatch representa un impacto muy notable para el avance del conocimiento científico de los factores relacionados con la distribución de la biodiversidad. Sin embargo, el mayor beneficio de la constitución de esta infraestructura se dará a futuro, con su contribución a la gestión medioambiental, al permitir identificar las zonas de mayor biodiversidad, determinar las áreas donde hay que incidir especialmente porque suponen los límites de la distribución de determinadas especies o poblaciones y, sobre todo, predecir los impactos antrópicos sobre la biodiversidad a las escalas local, regional, nacional o global. Al mismo tiempo, servirá para identificar y predecir la distribución de diferentes vectores de enfermedades o plagas, y favorecer así a su control.

La participación en esta e-infraestructura ofrece una oportunidad para estructurar la comunidad científica española relacionada con el estudio de la biodiversidad, así como para identificar las necesidades de los gestores medioambientales.

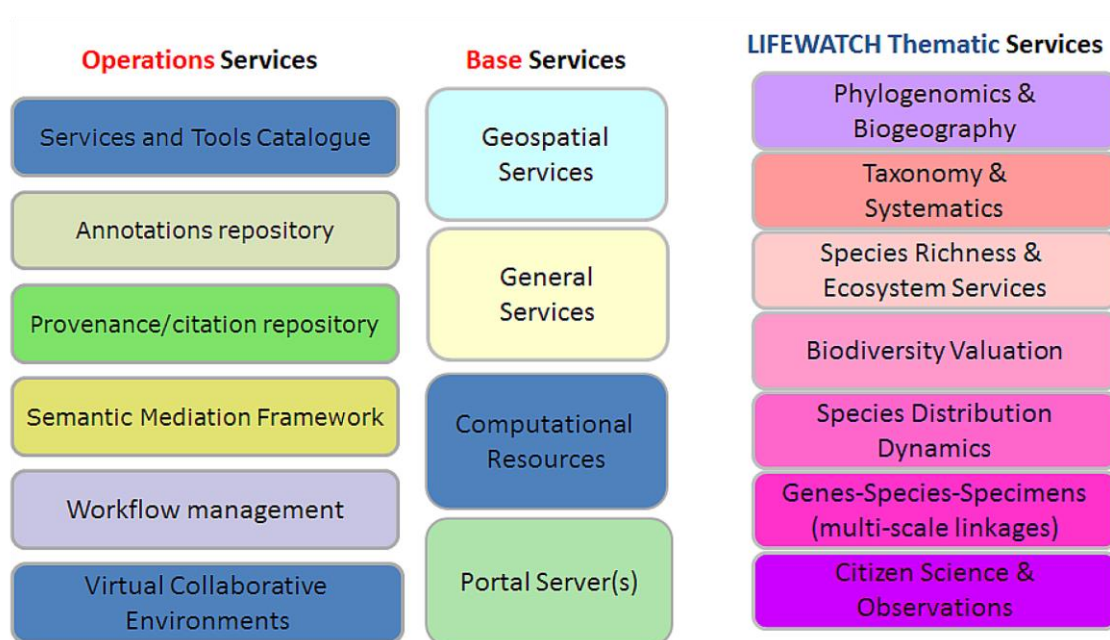


Figura 78. Ejemplos de los principales servicios que proporcionará la plataforma *LifeWatch*.
Fuente: LifeWatch, 2015.

Esto permite reforzar la posición de España en las áreas científicas de ecología y medio ambiente y planta-animal, donde ya ejerce un importante liderazgo. Al mismo tiempo, hace posible rentabilizar en mayor medida la participación española en diversas redes y organismos internacionales relacionados con la observación y la compilación de información de la biodiversidad. España aporta su conocimiento y experiencia respecto a e-infraestructuras y potencia de cálculo, por lo que LifeWatch se convierte una excelente plataforma para las empresas españolas relacionadas con las tecnologías de la información y el desarrollo de *software*.

LifeWatch tiene un contexto europeo y, por lo tanto, interactúa con otras infraestructuras e iniciativas paneuropeas (LifeWatch, 2015). La finalidad de estas interacciones no es otra que la de recoger y analizar conjuntos de datos enmarcados dentro de distintos ámbitos. Además, LifeWatch tiene como finalidad poder interactuar con otras e-infraestructuras regionales o globales (GBIF, GEO BON o GEOSS, por ejemplo).

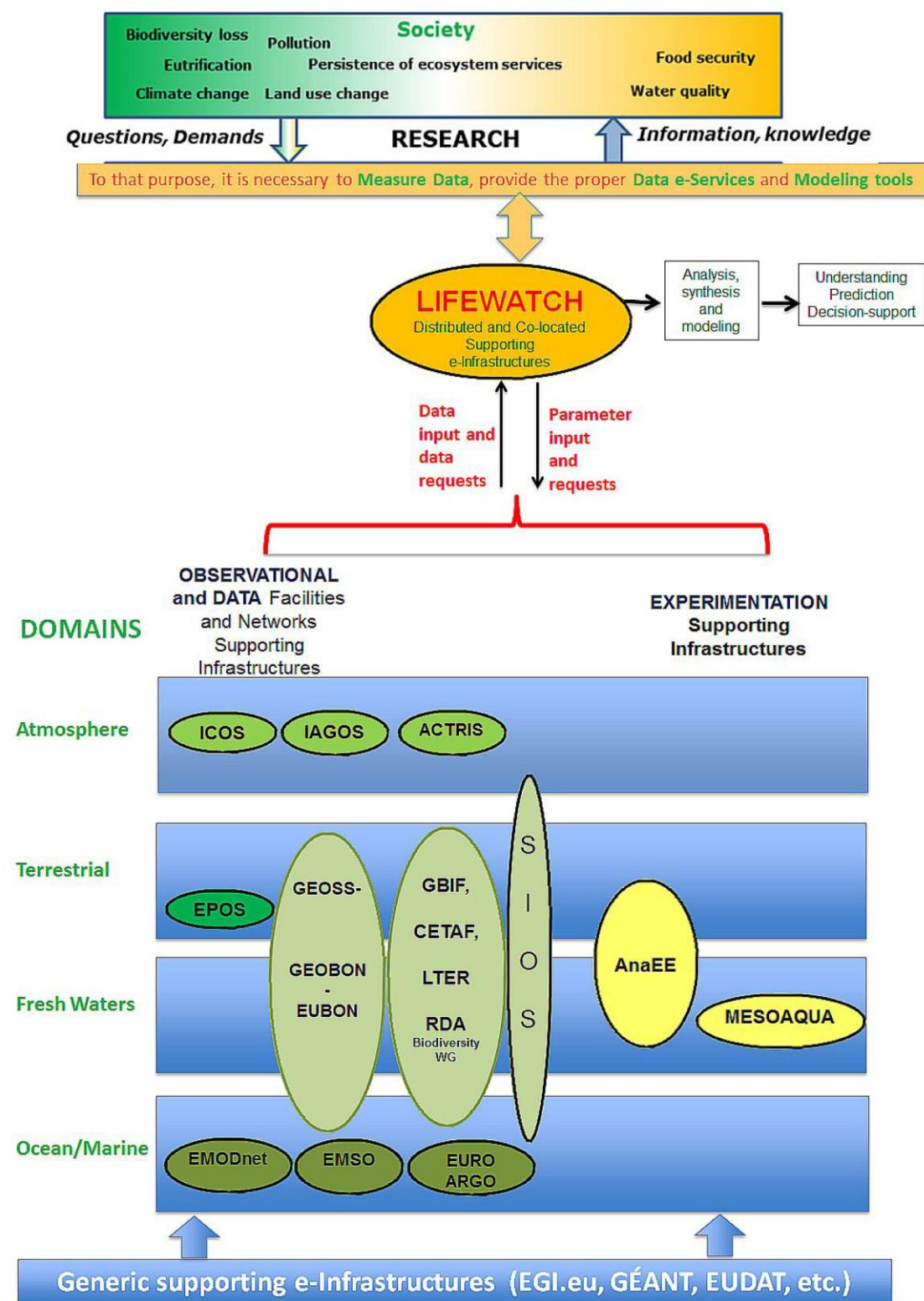


Figura 79. Relación de LifeWatch con otras infraestructuras e iniciativas pan-europeas. Fuente: LifeWatch, 2015.

Sin lugar a dudas, LifeWatch está llamada a convertirse en la infraestructura por excelencia no solo de datos de biodiversidad, sino también de análisis, incluso predictivo, de la interrelación de los datos abióticos (químicos, físicos, económicos,

etc.) con los de biodiversidad para crear productos (típicamente, programas *software* o de *middleware*, basados en las redes europeas de supercomputación como soporte *hardware*), es decir, aplicaciones informáticas y, especialmente, *virtual research environments* (VRE) que permitan modelizar y predecir el estado de la biodiversidad, sus amenazas y las posibles soluciones.

2. DOS EJEMPLOS DE ESTADOS UNIDOS

En Estados Unidos, dos modelos distintos precedieron al modelo europeo, aún en vías de construcción: por un lado, la Data Observation Network for Earth (DataONE) y, por otro, aunque añadiendo la función de repositorio mediante una red de sensores, la National Ecological Observatory Network (NEON).

2.1. El modelo de Data Observation Network for Earth (DataONE)

Este modelo se recoge en Allard (2012) y se explica como una colaboración (DataONE, s. f., What is DataONE) entre varias organizaciones asociadas, financiada por la National Science Foundation⁷⁷ estadounidense con arreglo al programa DataNET y se creó a través de un acuerdo de cooperación con la subvención número OCI-0830944. Recibió apoyo, adicionalmente, de otros organismos como la NSF Computer and Information Science Engineering Directorate (CISE) Pathways Computational Sustainability, NSF INTEROP Programs, Leon Levy Foundation, National Aeronautics and Space Administration (NASA), Microsoft Research & Gordon and Betty Moore Foundation, US Geological Survey (USGS) y National Biological Information Infrastructure (NBII).

A diferencia de Europa, en Estados Unidos el sistema de gestión de este tipo de e-infraestructuras se hace mediante la creación de estructuras de colaboración financiadas por distintas organizaciones asociadas (National Science Foundation, National Institute of Health, etc.). Supuso la fundación de una nueva e innovadora ciencia medioambiental articulada con un marco de trabajo distribuido y una ciberinfraestructura que cumplía con las necesidades tanto de la ciencia como de la sociedad en lo que respecta al acceso a datos bien descritos y descubiertos de la Tierra de forma abierta, duradera en el tiempo, consistente y segura.

Nació con la misión de garantizar la conservación, acceso, uso y reutilización de datos científicos multiescala, multidisciplinarios y multinacionales a través de tres elementos de la ciberinfraestructura primarios, una amplia educación y un ambicioso

⁷⁷ Las menciones a las organizaciones americanas dejan literales en este trabajo para que puedan encontrarse con facilidad en caso de necesidad de consultar en Internet.

programa. Se acaba de decidir el traspaso de su entramado institucional y operativo a la Universidad de California dentro de unos años.

2.1.1. La estructura de DataONE

Se construyó como una plataforma integradora para la investigación biológica y medioambiental (Michener *et al.*, 2012) que permitía desarrollar nueva ciencia y crear conocimiento mediante un acceso abierto, persistente, robusto y seguro a datos bien descritos y de fácil acceso en relación con la vida en la Tierra y el entorno que la sostiene. DataONE se estructuró como una red colaborativa (Sayogo y Pardo, 2011).

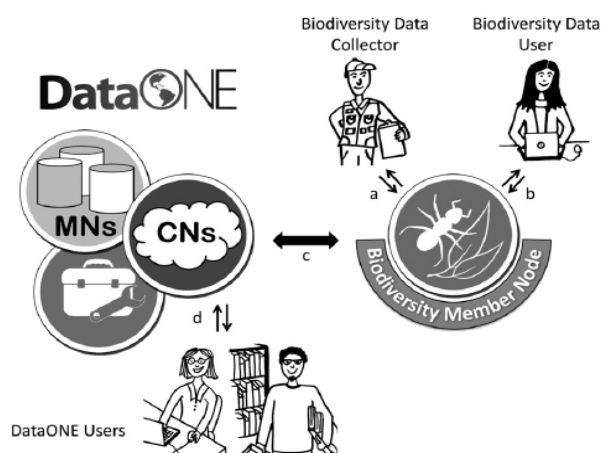


Figura 80. Resumen de los elementos de la ciberinfraestructura de DataONE que ilustran el valor que aporta a los usuarios asociados a un nodo miembro de biodiversidad. Colectores de datos de biodiversidad (a) depositar los datos en un repositorio de biodiversidad miembro de DataONE. Usuarios de datos de biodiversidad (b) pueden acceder a dichos datos directamente a través del repositorio de biodiversidad. La media luna alrededor del repositorio representa la pila de *software* que permite a los nodos miembros que el repositorio funcione. Esta pila de *software* la instalan los responsables de la ciberinfraestructura de DataONE. Los usuarios de biodiversidad pueden seguir accediendo a los datos tal como lo hacían antes (b); sin embargo, a través de la vinculación a la ciberinfraestructura de DataONE (c), dichos datos están accesibles a una comunidad más amplia de usuarios de DataONE (d) a través de la búsqueda y recuperación habilitados. Fuente: Michener *et al.*, 2012.

- **Nodos de coordinación:**

Se estructuró en tres nodos de coordinación con amplios servicios de red para potenciar la interoperabilidad de los miembros de los nodos y como soporte de los servicios de indexación y réplica de la información. Los nodos de coordinación proporcionaron una réplica de los catálogos de los nodos de los miembros y facilitan a los científicos el acceso a los datos con independencia de su lugar de residencia. Permitieron, asimismo, que los datos de los repositorios y los servicios que prestaban fueran más ampliamente accesibles para la comunidad internacional.

Los nodos de coordinación están situados en la Universidad de Nuevo México, en la Universidad Santa Bárbara de California y en la Universidad de Tennessee (en colaboración con el Laboratorio Nacional Oak Ridge) (DataONE, s. f., Organization).

- Nodos miembros:

Se creó integrando una red de centros de datos, una red de científicos y de organizaciones distribuidas geográficamente. Estas organizaciones podían exponer sus datos dentro de la red usando la interfaz de servicio de los nodos miembros. Además de los datos científicos, los nodos miembros podían proporcionar recursos informáticos o servicios adicionales como podía ser la réplica de datos de la comunidad de DataONE (DataONE, s. f., Organization).

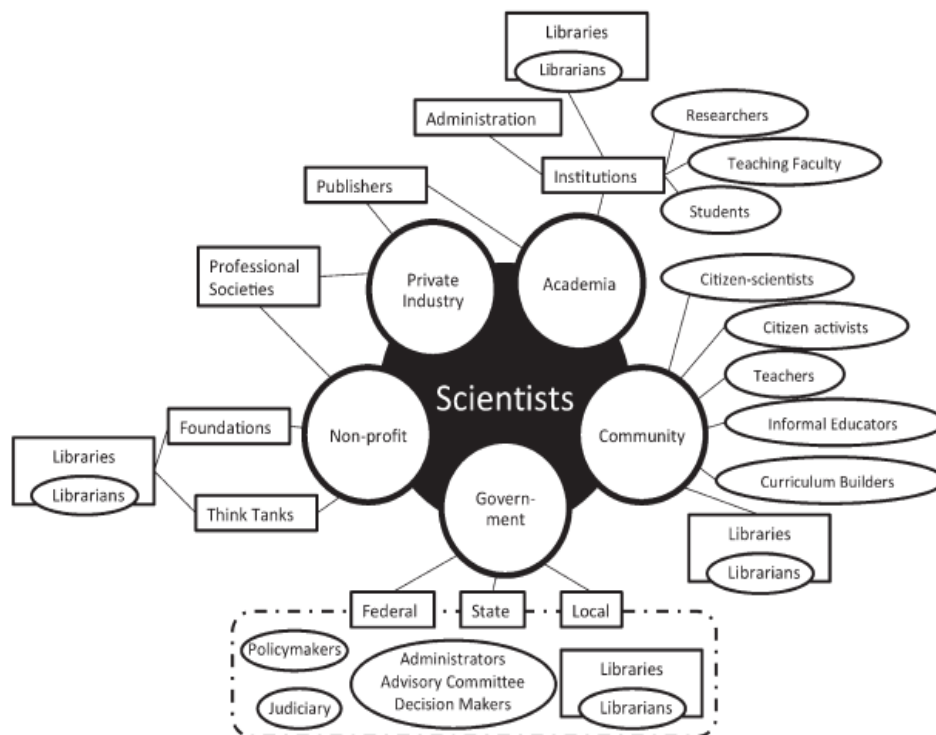


Figura 81. Relaciones entre las comunidades de stakeholders: los científicos son los stakeholders primarios y los círculos representan los entornos de investigación científica, los stakeholders secundarios están asociados con cada entorno de investigación científica e incluye organizaciones (cajas) e individuos (figuras ovaladas), las cajas con líneas discontinuas se refieren a los stakeholders asociados con cada nivel de gobierno. Fuente: Michener *et al.*, 2012.

La estructura aparecía descrita en el sitio web de la forma que se muestra en el siguiente gráfico.

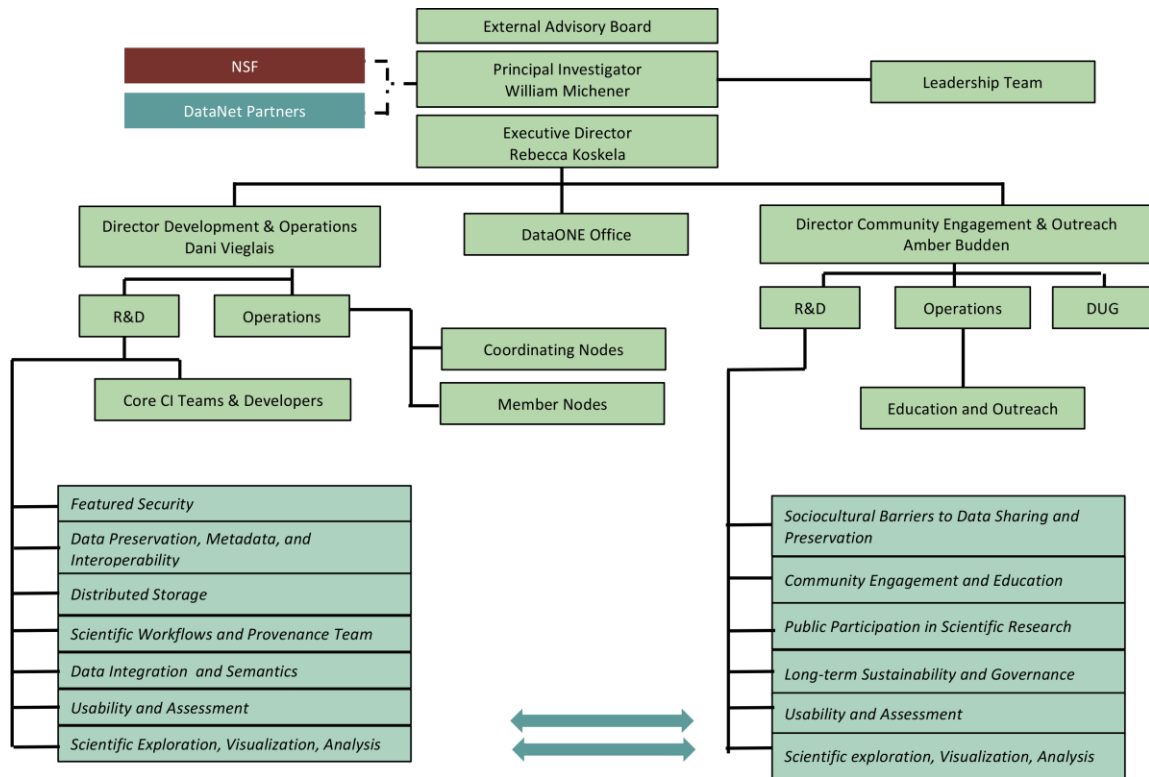


Figura 82. Organización de DataONE. Fuente: DataONE, s. f.

Ahora se ha modificado y es la siguiente que se muestra a continuación.

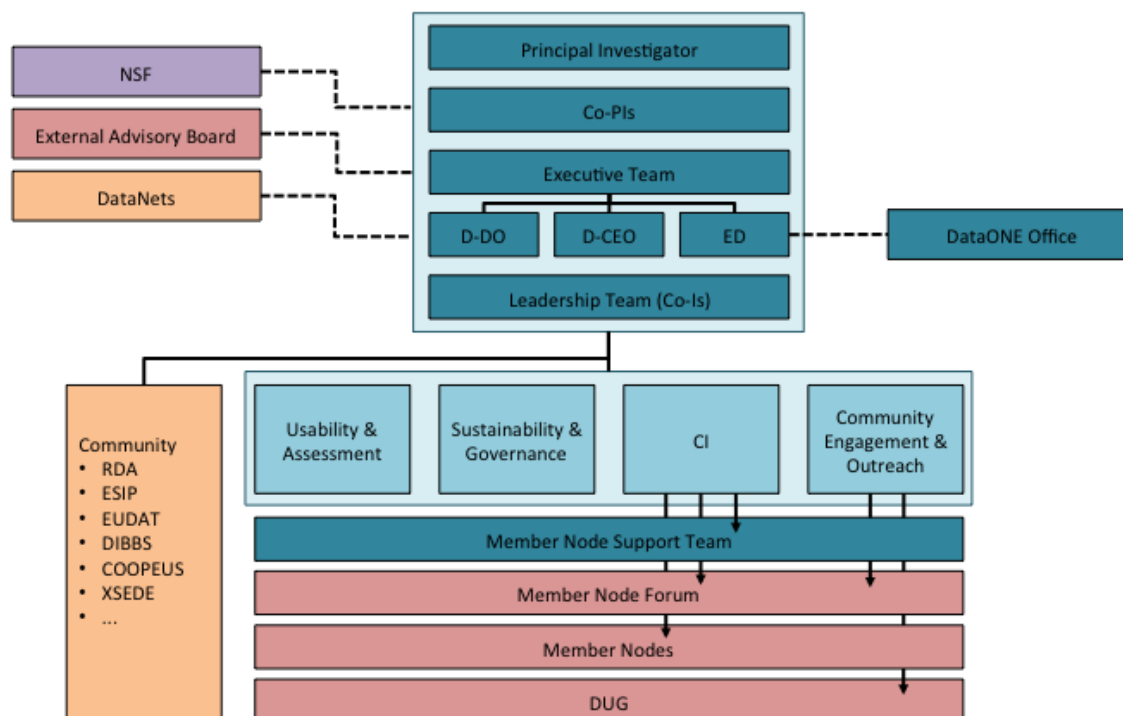


Figura 83. Estructura de DataONE. Fuente: DataONE, s. f..

El gráfico describe la estructura de relaciones organizativas entre varios grupos dentro de DataONE. Estas relaciones no son exclusivas y existen múltiples oportunidades de colaborar e integrar una mayor fortaleza.

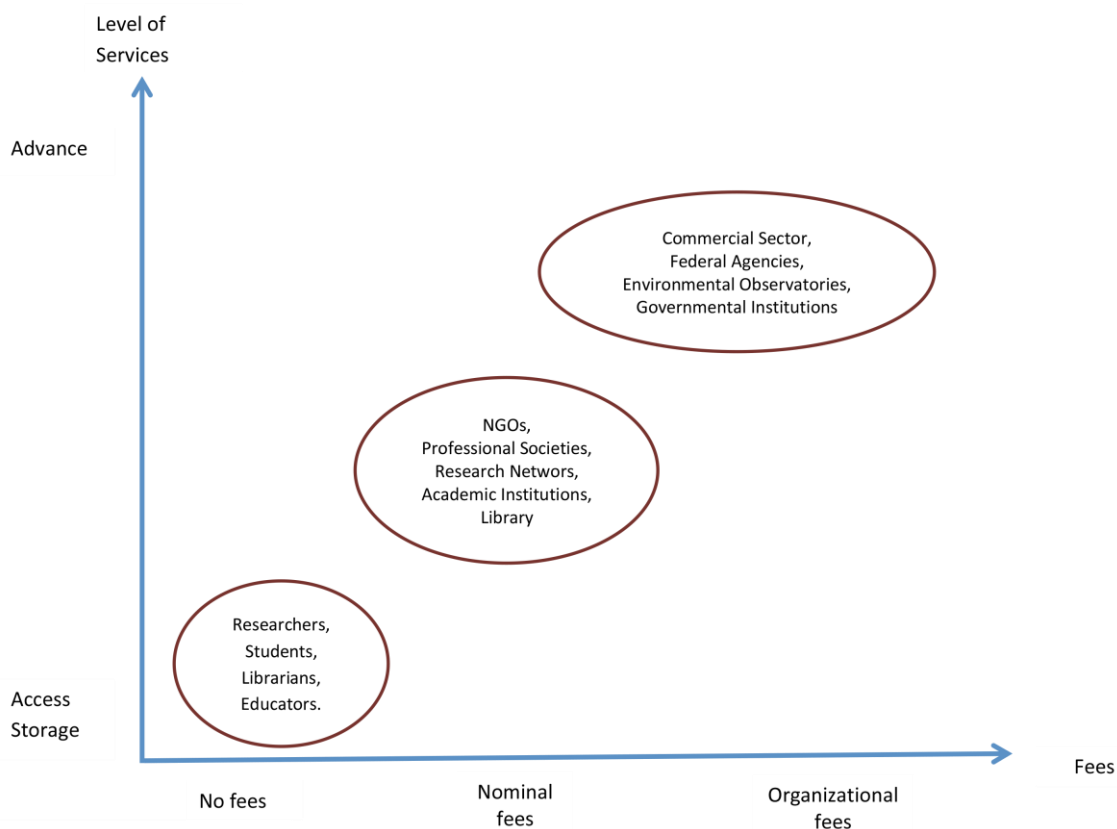


Figura 84. DataONE Membership Model. Fuente: Adaptada de DataONE, s.f.

La nueva era de grandes retos científicos y de posibilidades de becas ofrece un significativo potencial para el avance del nivel de conocimiento, para la transformación de la Academia y para la generación de beneficios para la sociedad (Michener, 2012). Tres acciones concretas pueden acelerar esta transición.

En primer lugar, es necesario promover este cambio integrando una ciencia intensiva en datos multidisciplinar y colaborativa. Esto requiere hacer *lobby* para lograr los fondos necesarios y dar apoyo y reconocimiento a los individuos que elijan sumarse a equipos para abordar los problemas ligados al gran reto. En segundo lugar, es necesario formar a las futuras generaciones de científicos, y no solo a aquellos que se dediquen a las ciencias computacionales e informáticas, inculcándoles conocimientos de la informática en sus currículos. En tercer lugar, es necesario abogar por un cambio dirigido a lograr un desglose de los datos académicos y una ruptura de los silos adecuadamente subvencionados, para lograr equipos multidisciplinarios de científicos preparados para abordar los grandes retos.

La arquitectura técnica de DataONE (Gunia y Sandursky, 2010) está diseñada para permitir a los científicos crear y depositar los datos descritos por cualquiera de las docenas de estándares de metadatos y formatos de archivos que, en la actualidad, se usan con carácter general, en las distintas disciplinas científicas. Cada uno de estos

estándares descriptivos de metadatos puede incluir distintas cantidades de metadatos estructurales, técnicos, de preservación y de administración.

DataONE ha definido la primera iteración de su sistema de metadatos: una lengua franca, añadida como capa de metadatos, para dar soporte al hallazgo, uso, reutilización y gestión de los contenidos digitales de DataONE. En resumen, la idea es disponer de una estructura que garantice la preservación de los datos a largo plazo.

DataONE fue diseñada como una infraestructura que facilita la preservación de los datos y su reutilización en investigación, centrada al principio en ciencias biológicas y del medioambiente (Michener *et al.*, 2012). Es una red de datos federada construida para mejorar el acceso a los datos de la vida en la tierra y del entorno que la sostiene así como para apoyar a la ciencia (Michener, 2011) del siguiente modo:

1. Extrayendo los datos científicos relevantes y las comunidades políticas.
2. Facilitando un almacenamiento fácil, seguro y duradero de los datos.
3. Diseminando herramientas integradas y fáciles de usar para encontrar los datos, analizarlos, visualizarlos y para adoptar las decisiones.

Fue diseñada y construida, como señala Michener (2011), para proporcionar una base para la investigación innovadora medioambiental que resolviese cuestiones relevantes para la sociedad y para la ciencia. El cumplimiento de su misión está ligado a cinco actividades cruciales:

1. *Descubrimiento y consulta*: permitir a través de un único sitio el descubrimiento y la consulta multiescala, multidisciplinar y multinacional.
2. *Integración de los datos y síntesis*: prestar ayuda con el desarrollo de herramientas de transformación que modelan nuestra comprensión de los procesos de la Tierra partiendo de un ámbito local hasta escalas globales.
3. *Formación y capacitación*: proporcionar las habilidades fundamentales que permitan potenciar la investigación científica (por ejemplo: formación en gestión de datos, compartir las denominadas *best practices* o buenas prácticas, descubrimiento de herramientas).
4. *Construir comunidad*: combinar el conocimiento y los recursos de distintas comunidades para educar, promover y apoyar la gestión fiable de los datos científicos.
5. *Compartir datos*: proporcionar incentivos y e-infraestructuras para compartir datos de investigadores subvencionados por la academia.

Se configuró como una colaboración multiinstitucional, multinacional e interdisciplinaria de trabajo cuya finalidad era desarrollar una estructura organizativa que apoyase la información completa del ciclo de vida de los datos y herramientas biológicas, ecológicas y ambientales para que las pudieran utilizar los investigadores, educadores y el público en general.

Fue única en la medida en que se construyó partiendo de centros de datos existentes y ponderando las inversiones globales en preservación de datos científicos. Además, creó una red de datos federada orientada a dar soluciones de interoperabilidad, que proporciona herramientas y servicios para la creación de nueva ciencia y de nuevo conocimiento. Facilita la creación de comunidades de práctica dentro de la ciberinfraestructura de DataONE (Michener *et al.*, 2012).

Su finalidad primordial era permitir la investigación biológica y ambiental intensiva en datos mediante una ciberinfraestructura que podía utilizarse como herramienta de la nueva ciencia. La clave estaba en que los datos debían ser robustos, accesibles y seguros, por eso era crucial la gestión de datos, tanto desde el punto de vista técnico como desde el sociocultural.

Abordaba tres problemas. En primer lugar, daba soporte al estudio de problemas ambientales complejos, por ejemplo: el cambio climático. Las cuestiones ambientales representan sistemas adaptativos complejos que pueden afectar a muchas disciplinas diferentes y requerir conocimientos de diferentes áreas académicas.

El segundo problema fue la falta de uso de formas compatibles de almacenar datos. Existen retos adicionales en relación con los datos, como podía ser la pérdida de datos (desastre natural, la obsolescencia del formato, los datos huérfanos), las fuentes de datos dispersas, diluvio de datos (el flujo de datos cada vez más heterogéneos), las prácticas deficientes de datos y la durabilidad de los datos.

El tercer problema fue la necesidad de abordar un problema global con una perspectiva global. No es suficiente abordar estas cuestiones de forma parcial. Diseñada para hacer frente a la necesidad de acceso a datos, seguro y sólido, fundamental para que los esfuerzos de investigación sean productivos y para que puedan elaborarse políticas sobre cuestiones ambientales, DataONE perseguía los siguientes objetivos:

1. El acceso coordinado a la base de datos actual utilizando la ciberinfraestructura disponible.
2. La creación de una nueva infraestructura cibernética mundial con datos tanto biológicos como ambientales, procedentes de distintas fuentes (por

ejemplo: redes de investigación, observatorios de medio ambiente, científicos individuales y ciudadanos).

3. Cambiar la cultura y las instituciones de la ciencia a través de la nueva ciberinfraestructura práctica, proporcionando educación y formación, participación de los ciudadanos en la ciencia, y la construcción de comunidades globales de práctica. Esto lleva a la definición de su misión, apoyar la ciencia a través de tres áreas principales: proporcionar un conjunto de herramientas para el descubrimiento de datos, análisis, visualización y la toma de decisiones, prestar servicios de almacenamiento de datos de forma fácil, segura y persistente y facilitar la participación de la comunidad de científicos, especialistas de datos y reguladores.

El modelo de DataONE no es el único existente en Estados Unidos, como se verá en el punto siguiente.

2.2. El ejemplo de la National Ecological Observatory Network (NEON)

Fue un proyecto que tuvo su origen en las potestades de gasto del propio Gobierno federal y, en concreto, de la oficina del presidente. Patrocinado por la National Science Foundation, se creó para recoger y proporcionar datos en abierto que caracterizasen y cuantificasen los rápidos y complejos cambios ecológicos que tienen lugar dentro del territorio de Estados Unidos. Los datos globales, la extensión espacial y la tecnología de teleobservación que proporcionaba NEON hizo posible que una comunidad más amplia de usuarios pudiera abordar cuestiones nuevas a escalas no accesibles a generaciones anteriores de ecologistas (NEON, s. f.).

Fue un audaz esfuerzo por construir partiendo de los recientes progresos en muchos ámbitos con la finalidad de abrir nuevos horizontes a la ciencia de la ecología a gran escala. Dio importancia, explícitamente, a cuestiones relacionadas con los grandes desafíos de la ciencia ambiental, relevantes para las grandes regiones de los Estados Unidos, y a los que no puede hacerse frente con los enfoques ecológicos tradicionales (Field *et al.*, 2006).

El observatorio fue específicamente diseñado para cuestiones científicas clave sobre las interacciones de los ecosistemas, el clima y el uso de la Tierra (Field *et al.*, 2006). El planteamiento conceptual de estos últimos conduce a plantearse las siguientes cuestiones (Keller *et al.*, 2008):

1. ¿Cómo responderán los ecosistemas (de Estados Unidos) y sus componentes a los cambios en forzamientos naturales e inducidos por el hombre, como puede ser el clima, el uso de la tierra y las especies invasoras,

en todo un rango de escalas espaciales y temporales? ¿Cuál es el ritmo y la pauta de dichas respuestas?

2. ¿De qué manera interactúan las respuestas internas y la retroalimentación de la biogeoquímica, la biodiversidad, la hidroecología y la estructura y funciones bióticas con los cambios en el clima, el uso de la Tierra y las especies invasoras? ¿Cómo varían estas realimentaciones en el contexto ecológico y las escalas espaciales y temporales?

NEON nació con la idea de operar durante treinta años y recabar datos a largo plazo sobre la respuesta de la ecología a los cambios y, también, sobre las retroalimentaciones con la geosfera, la biosfera y la atmósfera (Kampe *et al.*, 2010).

2.2.1. La estructura de NEON

Según apuntan Field *et al.* (2006) se basó en un diseño estratificado combinando una amplia cobertura de Estados Unidos por parte de agencias de colaboración que gestionaban plataformas espaciales y de aeronaves con un despliegue más intenso de infraestructuras en áreas representativas cuidadosamente seleccionadas, puntos críticos para comprender, en particular, los motores concretos y las respuestas al cambio, así como otros lugares de especial interés.

Además, el observatorio ofreció la posibilidad de desplegar redes de experimentos en laboratorio para explorar partes del dominio del espacio-tiempo inaccesibles con otros enfoques.

NEON vinculó diversas tecnologías y enfoques en un marco coordinado que proporcionaban mediciones estandarizadas y de alta calidad durante un periodo de tiempo prolongado en ubicaciones y en diferentes grados que permitían el acceso a una amplia gama de ecosistemas, procesos y retroalimentación a gran escala.

Ha realizado importantes inversiones en siete aspectos interrelacionados de su programa de investigación.

1. Veinte sitios nucleares fuertemente instrumentados, localizados en áreas silvestres seleccionadas para abarcar la gama de las principales zonas climáticas y ecosistemas de Estados Unidos.
2. Plataformas aéreas de observación con instrumentos de teledetección para proporcionar información regional para escalamiento y extrapolación desde sitios.
3. Lugares en distintos grados, atendidos por sistemas móviles o reubicables, localizados para facilitar la comprensión a través de observaciones de ecosistemas expuestos a diferencias a largo plazo en factores clave

ambientales o de dimensión humana, tales como elevación, precipitación, uso de la Tierra, tiempo transcurrido desde el disturbio y ubicación dentro de una cuenca hidrográfica importante.

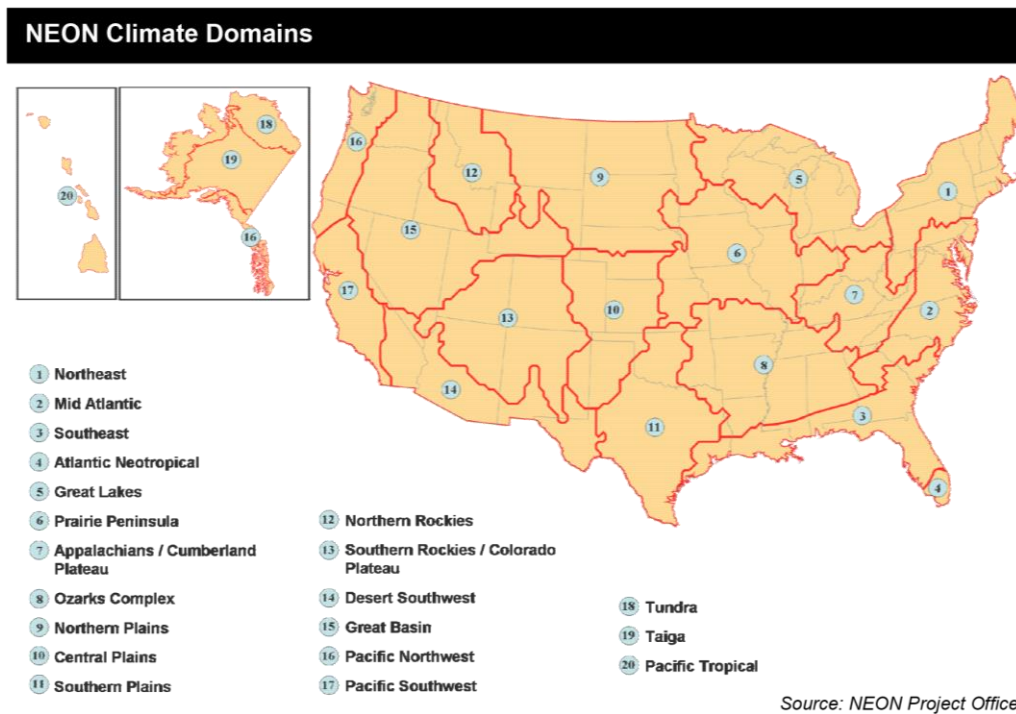


Figura 85. Neon Climate Domains. Fuente: NEON, 2006.

4. Sitios de oportunidad atendidos por sistemas móviles seleccionados para permitir el estudio detallado del proceso, tendencia o intervención, como la recuperación de un incendio forestal importante, la respuesta a una invasión de especies, la conversión a la agricultura, el abandono de la agricultura o la rehabilitación de una cuenca hidrográfica urbana.
5. Experimentos a escala nacional, sometidos a experimentos coordinados (como el calentamiento, alteraciones de las precipitaciones y reducción de la biodiversidad) para evaluar los controles a gran escala sobre las respuestas de los ecosistemas, en particular, en procesos en los que los gradientes espaciales no son informativos.
6. Ciberinfraestructuras para dar soporte a las diversas observaciones y modelos, y gestionar los datos y resultados que generan, incluyendo redes de sensores autónomos y sistemas de seguimiento de animales de alta resolución.
7. Programas de educación y divulgación que preparen a la comunidad científica para maximizar el uso de NEON, promoviendo la ciencia mediante la ampliación y diversificación de la comunidad de investigación y

dando soporte al público para usar y beneficiarse de NEON para lograr una mejor comprensión para abordar eficazmente cuestiones ecológicas críticas.

NEON tiene vocación de servir de puente entre la diversidad de escalas, mediante el análisis de procesos, interacciones y respuestas, incluidas aquellas en las que intervengan el transporte y la conectividad.

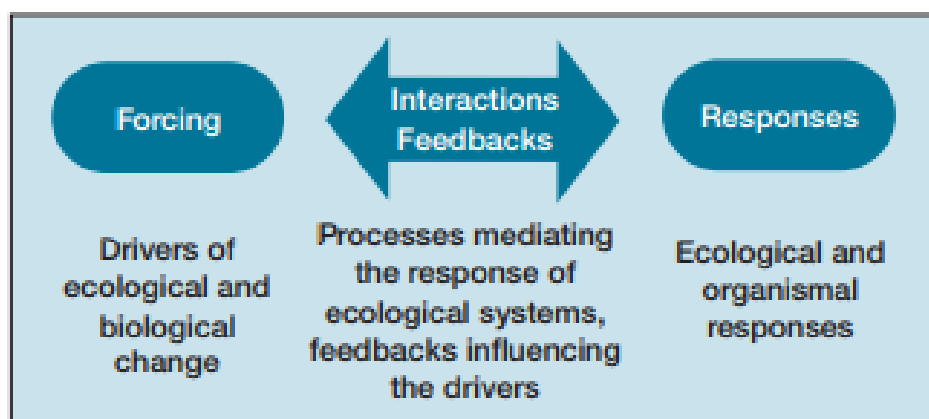


Figura 86. NEON difiere de otras redes de monitorización medioambiental porque, en su diseño, integra procesos, interacciones y respuestas. Fuente: Keller *et al.*, 2008.

La estrategia de observación y el proceso de selección de sitios de NEON se han ido haciendo con base en un muestreo sistemático en escalas más grandes de la variabilidad ecológica proporcionando una base para los análisis de que puedan escalar a todo el país. Ha dividido el continente en regímenes ecoclimáticos llamados «dominios». Estados Unidos y Puerto Rico tienen diecisiete dominios, y Alaska y Hawái tres más, lo que da un total de veinte. La red incluye elementos estables, fijos (sitios forestales clave), lugares reubicables y laboratorios móviles. Los dominios de NEON han sido elegidos como elementos de una estrategia de observación de escala continental, basada en la teoría e informada por una amplia gama de conjuntos de datos y enfoques estadísticos (Schimel *et al.*, 2007).

El diseño de muestreo espacial de NEON está basado en atributos climáticos, edáficos y topográficos, delimitados en dominios que usan datos de clima y de las propiedades del suelo. Los dominios se mapearon de acuerdo a variables físicas que efectivamente capturaron aspectos biológicos clave de la ecología estadounidense. Los dominios NEON se basaron en un procedimiento estadístico multifactorial explícito transparente y repetible, y proporcionan una estratificación nacional que muestra la mejor forma de desplegar los veinte sitios principales para maximizar la cobertura, coherencia y representatividad de la red. Los dominios también se pueden utilizar para extrapolar las mediciones realizadas en los sitios, facilitando el escalado. Este plan de muestreo, basado en veinte dominios, convirtió a NEON en

la red ecológica más grande y completa diseñada estadísticamente antes de que esta fuera implementada (Schimel *et al.*, 2007).

Su diseño puede interpretarse como una jerarquía de restricciones, cuando, en realidad, es frecuente pensar en el clima como una variable independiente en relación con la biología. En la escala continental las variables climáticas sufren restricciones derivadas de la latitud, la «continentalidad» y la orografía (la geografía física de las montañas). La influencia de las cuencas oceánicas que rodean al continente también se filtra a cada uno de los dominios, y afecta a la temperatura, las precipitaciones y a la variabilidad en el tiempo (por ejemplo: vía El Niño). Los sitios clave forestales abordan estas escalas más grandes (Schimel *et al.*, 2007).

Las pautas propias del uso de la Tierra, de la gestión, de las perturbaciones (como el fuego o las inundaciones) y de la recuperación se desarrollan dentro de los dominios, limitados por el entorno biofísico, la biota local y los efectos históricos. Las distintas pautas de uso y asentamiento de los recursos naturales tienden a evolucionar en cada dominio (por ejemplo: producción de madera en áreas boscosas y agricultura en regiones con abundante fertilidad y precipitación o riego). Estos diversos patrones de regímenes socioambientales permiten estudiar las interacciones de los ecosistemas con la dinámica humana. Los sitios de gradientes reubicables son muestras de estas escalas más pequeñas de variación (Schimel *et al.*, 2007).

Los ecologistas han explotado principalmente la variabilidad natural a lo largo de gradientes en los que solo varía un factor. Los dominios de NEON no representan sitios elegidos de acuerdo a un único factor; en su lugar, muestrean regímenes en los que el clima, la biota, los suelos y las prácticas de manejo de la Tierra varían juntos. Aunque este diseño no permite la inferencia estadística simple, proporciona diversas condiciones mediante las cuales se pueden abordar hipótesis, preguntas y modelos. Los modelos estadísticos modernos y de procesos son ahora lo bastante sofisticados para ser utilizados con este diseño de régimen complejo e interactivo (Schimel *et al.*, 2007).

Sigue existiendo una gran flexibilidad en los instrumentos y medidas a desplegar y en los futuros sitios de las instalaciones reubicables, de manera que el diseño actual pueda evolucionar a medida que la ciencia vaya madurando. El diseño actual de la red equilibra el amplio sistema de cobertura nacional con la relevancia de los sitios y subredes individuales con las cuestiones locales y regionales. Muchas decisiones siguen siendo difíciles, y NEON depende de los *inputs* de la comunidad (Schimel *et al.*, 2007).

La ecología está evolucionando rápidamente y transformándose, paulatinamente, en una ciencia más abierta, responsable, interdisciplinaria, colaborativa e intensiva en datos. El descubrimiento, la integración y el análisis de grandes cantidades de datos

heterogéneos son fundamentales para la misma, ya que los investigadores abordan cuestiones complejas a escalas tan amplias como puede ser desde el gen hasta la biosfera (Michener y Jones, 2012).

Nuevas redes de observatorios en Estados Unidos, tales como NEON, proporcionan plataformas de investigación que permiten a los científicos examinar fenómenos a través de diversos tipos de ecosistemas. Se ha postulado que la ciencia intensiva en datos representa el cuarto paradigma científico que sigue los enfoques empíricos (es decir, descripción de los fenómenos naturales), teóricos (por ejemplo: modelización y generalización) y computacionales (por ejemplo, simulación), y comprende un enfoque para unificar la teoría y la simulación (Michener y Jones, 2012).

3. OTROS MODELOS CONTINENTALES

Además de los modelos de Europa y de Estados Unidos, analizados en detalle, existen otros modelos que se han probado a escala continental que merece la pena enunciar y describir brevemente. En primer lugar, para mostrar otras iniciativas distintas nacionales y, en segundo lugar, porque estas e-infraestructuras formarán parte de los ejemplos de *roadmaps* realizados en la Unión Europea como se verá después en el punto relacionado con los *roadmaps*.

3.1. South African National Biodiversity Institute

El South African National Biodiversity Institute (SANBI) es el líder y coordinador de la investigación y del seguimiento y el responsable de informar sobre el estado de la biodiversidad en África del Sur. El punto fundamental de SANBI es suministrar información de biodiversidad. Lo hace con varias bases de datos y otras fuentes que han desarrollado SANBI y sus socios.

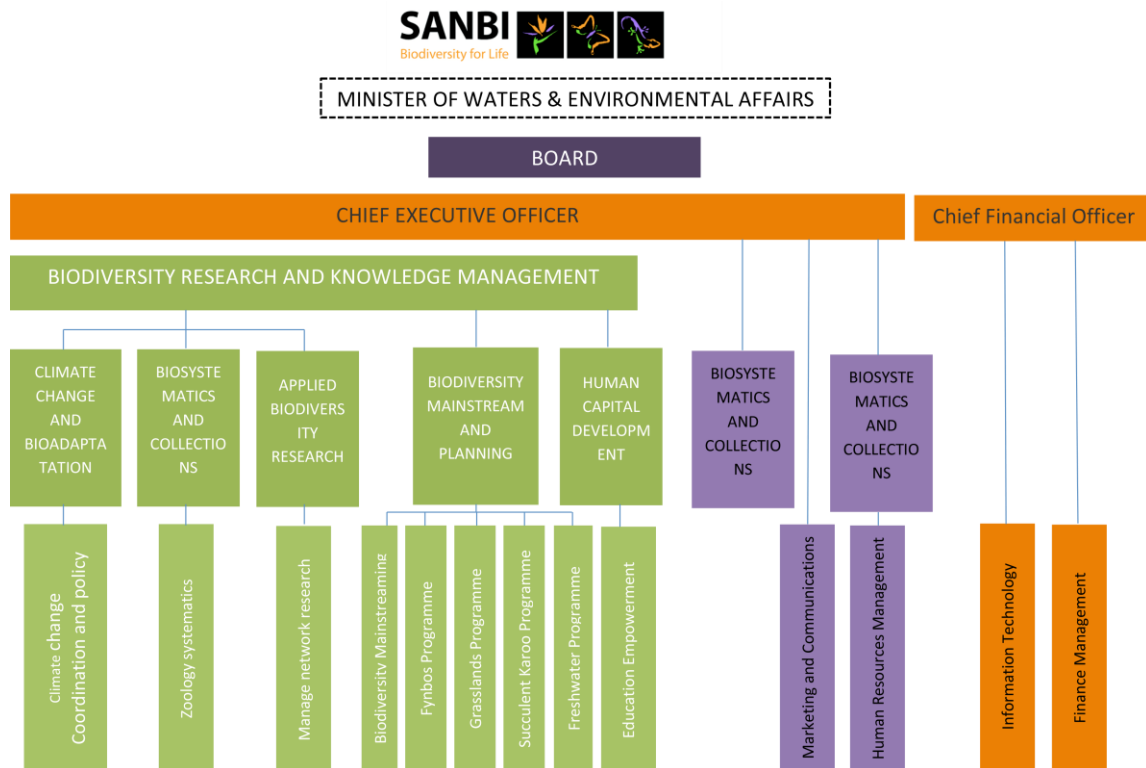


Figura 87. Estructura SANBI. Fuente: Alonso García, 2014.

Sudáfrica es uno de los países con una mayor biodiversidad en el mundo (después de Indonesia y Brasil). Es un país rodeado por dos océanos. Aunque en superficie representa menos del 2% de la superficie mundial tiene el 10% de las plantas que existen en la Tierra, un 7% de los reptiles, pájaros y mamíferos y un 15% de las especies marinas que viven en las costas. Además existen nueve biomas (paisajes con vegetaciones únicas), de los cuales tres han sido declarados regiones clave para la biodiversidad mundial. Por lo tanto, la biodiversidad es uno de los activos principales de Sudáfrica.

Esta organización fue creada en 2004 con una ley dictada por el entonces presidente de Sudáfrica Thabo Mbeki: la National Environmental Management: Biodiversity Act 10 of 2004. Por consiguiente, es una organización pública. Está integrada por una serie de institutos y organizaciones y está extendiendo su ámbito de influencia convirtiéndose en la referencia de los países de del sur de África.



Figura 88. Trabajo de Sanbi. Fuente: Sanbi, 2016.

Tiene, por consiguiente, un mandato legal para gestionar la información de biodiversidad. Una gestión eficiente de la información sustenta la investigación, la toma de decisiones, el asesoramiento de las políticas y la supervisión.

La Dirección de Información y Planificación de la Biodiversidad (Biodiversity Information and Planning Directorate), que resulta de la fusión de la Dirección de Gestión de Información de Biodiversidad (Biodiversity Information Management Directorate) y de la Dirección de Integración y Planificación de la Biodiversidad (Biodiversity Planning and Mainstreaming Directorate), presta servicios a SANBI. Es responsable la movilización y publicación de datos de biodiversidad para que los usuarios puedan acceder libremente a los mismos.

SANBI moviliza información primaria de biodiversidad para ponerla a disposición de los investigadores, de forma gratuita, en Internet y con la finalidad de que los que realizan las políticas, los gestores y los investigadores puedan tomar decisiones informadas que contribuyan al desarrollo sostenible de Sudáfrica.

Lo hace, fundamentalmente de cuatro maneras. En primer lugar, promueve el intercambio de datos e información mediante el establecimiento de unos estándares:

- Desarrollando y manteniendo de bases de metadatos.
- Estableciendo y manteniendo un portal web.

- Promoviendo las subvenciones relacionadas con la digitalización de datos.
- Dando subvenciones para interfaces de proveedores de datos.
- Facilitando el desarrollo de herramientas y aplicaciones para el uso de los datos.
- Promoviendo la coordinación e interconexión de actividades.

En segundo lugar, contribuye a la enseñanza, formación y desarrollo de la capacitación con el objeto de promover el acceso nacional a los datos. En tercer lugar, facilita la concienciación y divulgación de los beneficios y aplicaciones para los usuarios finales. Por último, asesora en relación con actividades nacionales e internacionales. Promueve además el acceso abierto a los datos, aunque existe una multiplicidad de limitaciones a ese acceso abierto.

3.2. Centro de Referencia em Informação Ambiental (CRIA), Brasil

CRIA es una asociación privada sin ánimo de lucro que persigue la diseminación de la información electrónica como herramienta en la organización de las comunidades científicas y tecnológicas del país.⁷⁸

Creada en el año 2000 como organización sin ánimo de lucro con la misión de contribuir al uso sostenible de la biodiversidad en Brasil, en 2002, el Ministerio de Justicia brasileño le concedió el estatus de organización civil de interés público para la sociedad.

Es la organización del desarrollo y mantenimiento de las redes de especies, actuando conjuntamente con varias colecciones biológicas responsables de los datos disponibles. Contribuye directamente a la conservación y utilización racional de la biodiversidad en Brasil.

En Brasil, CRIA agrega y disemina la información biológica de interés ambiental e industrial como una forma de organizar la comunidad científica y tecnológica del país dirigiéndola a la conservación y al uso sostenible de los recursos biológicos de Brasil.

El objetivo fundamental de CRIA es desarrollar la e-infraestructura brasileña de datos de biodiversidad. Además del esfuerzo continuado por mejorar la infraestructura de datos de biodiversidad, está ahora enfocada en el desarrollo y la

⁷⁸ Para más detalle puede consultarse <http://inct.florabrasil.net/participantes/instituicoes-associadas/centro-de-referencia-em-informacao-ambiental-cria>. (Consultado por última vez el 1 de abril de 2017).

implementación de aplicaciones y de servicios de último nivel para dar soporte a la disseminación y análisis de datos de biodiversidad. Colabora con organizaciones de ámbito nacional e internacional.

Sirve también de soporte para llevar a cabo la interoperabilidad (este punto se analizará con detalle en el apartado relacionado con la interoperabilidad) requerida para realizar el análisis de la biodiversidad usando la computación en la nube. El hecho de que CRIA se preocupe principalmente de la disseminación de información electrónica hace que haya tenido un gran éxito en la promoción del intercambio de datos y en la libre publicación y difusión de los datos de biodiversidad.

3.3. The Atlas of Living Australia (ALA)

Una de las mayores barreras en Australia para la investigación y gestión de la biodiversidad ha sido la fragmentación y la imposibilidad de acceder a algunos de los datos de biodiversidad. La información podía encontrarse en los herbarios, museos, universidades, organizaciones y departamentos del Gobierno.

Como consecuencia de ello, la obtención de los datos era costosa en tiempo y dinero y, además, resultaba en información incompleta. Por eso, el Gobierno australiano decidió construir una herramienta colaborativa con la finalidad de intercambiar y analizar la información de biodiversidad de Australia.

El ALA⁷⁹ es una e-infraestructura creada por el Gobierno australiano a través de Estrategia de Infraestructuras de Investigación Colaborativas (conocida por las siglas en inglés NCRIS, de National Collaborative Research Infrastructure Strategy) (Atlas of living Australia sharing biodiversity knowledge, 2016).

Es una asociación colaborativa de organizaciones que gestionan datos biológicos y con conocimiento en bioinformática, dentro de la que se encuentran museos, colecciones de datos biológicos, organizaciones de investigación, Gobiernos estatales y federales y gestores de recursos naturales.

Tiene como objetivo proporcionar una infraestructura basada en sitio web que permite capturar, agregar, gestionar, descubrir y analizar todo tipo de datos de biodiversidad e información asociada, mediante una serie de herramientas y capas espaciales que pueden usarse para investigación, industria, Gobierno y para la comunidad.

Contiene más de sesenta y siete millones de registros de especímenes, campos de observación y encuestas. Estos registros pueden enriquecerse con información

⁷⁹ Para más detalles, puede consultarse su página web en <http://www.ala.org.au>. (Consultado por última vez el 1 de abril de 2017).

adicional incluidos los datos moleculares, las fotografías, los mapas, grabaciones que tengan sentido y literatura. Contiene, por lo tanto, información de todas las especies conocidas de Australia, procedentes de un amplio rango de proveedores de datos: museos, herbarios, comunidades, departamentos del Gobierno, individuos y universidades.

La visión de esta infraestructura es liderar la transformación digital en lo que afecta al intercambio de conocimiento de biodiversidad dando soporte y haciendo posible que con los resultados procedentes de la investigación de alta calidad y de la innovación pueda hacerse frente a los retos tanto nacionales como globales.

Además, sirve de soporte para los *stakeholders* de investigación, descubrimientos de biodiversidad y documentación, seguimiento medioambiental e información, planificación de la conservación, actividades de bioseguridad, educación y ciencia ciudadana simultáneamente con compañías y organizaciones que sirven de contrapeso a la infraestructura abierta para crear y promover sus propios servicios y productos. Es una de las primeras e-infraestructuras creadas en Australia. Contribuye activamente a los avances que se están produciendo en gestión de los datos y es, además, el nodo australiano del GBIF.

Está basado en los principios de *open access*, analizados en el punto relacionado con la interoperabilidad. En resumen, pretende que los datos que se recojan una vez puedan ser de libre acceso y ser libremente reutilizables. Esta característica es particularmente importante en lo que afecta al concepto de datos públicos, aquellos que se recaban o producen como consecuencia de proyectos subvencionados por el Gobierno.

Promueve, además, un enfoque colaborativo y una comunidad de práctica en torno a la infraestructura y a los beneficios de interoperabilidad y de ahorro de costes. Los temas de interoperabilidad se abordarán en profundidad en el apartado relacionado con estas cuestiones.

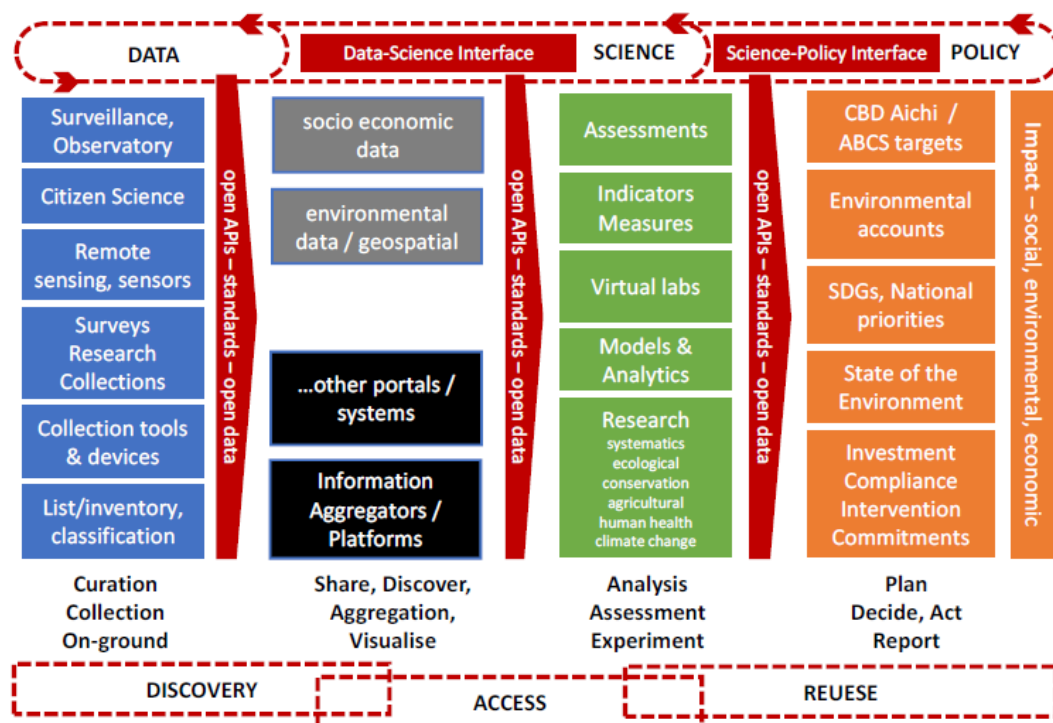


Figura 89. Implementación de la infraestructura de ALA. Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.

Overview of ALA infrastructure

The following diagrams provide an overview of the key components, tools and functions of the ALA infrastructure:

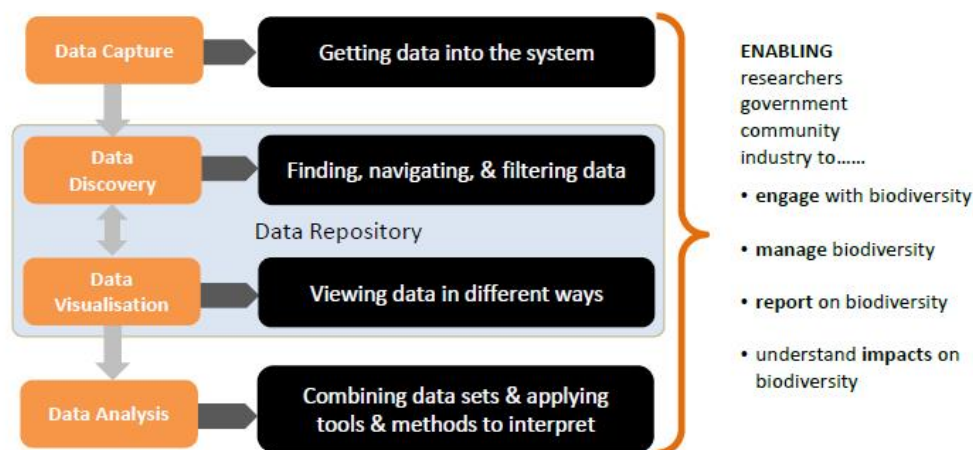


Figure 2 – ALA components and their usage

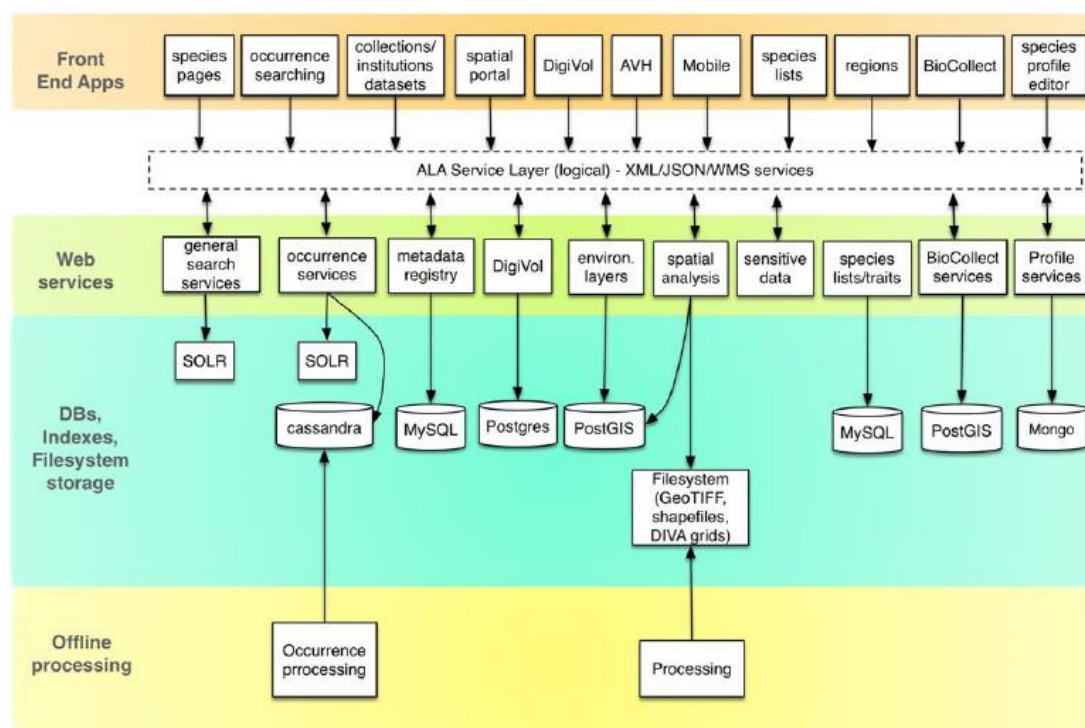


Figura 90. Infraestructura de ALA. Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.

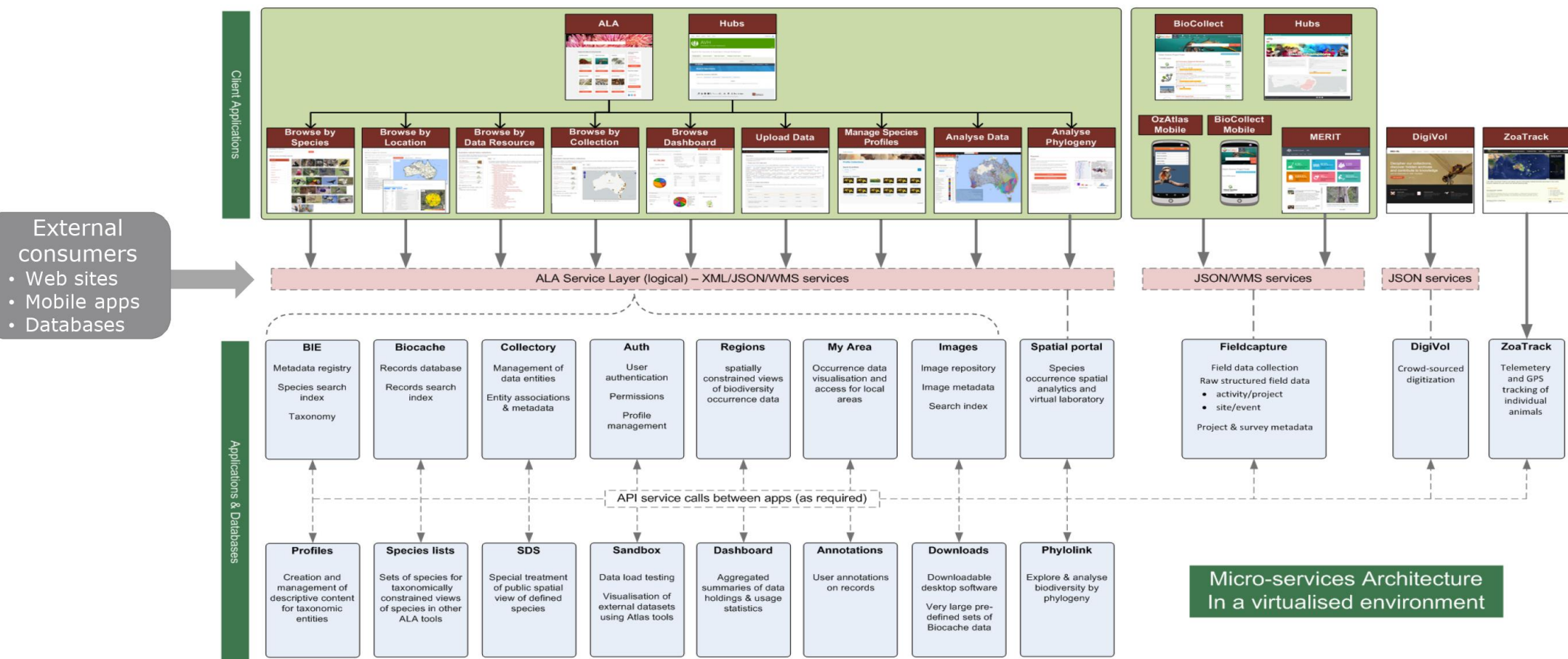


Figura 91. Modelo de arquitectura de ALA. Fuente: Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.

3.4. Academia de Ciencias de China (CAS Chinese Academy of Science)

Es una organización gubernamental, no una infraestructura en sí misma. Se fundó en Pekín en 1949. Es la institución académica nacional de más alto nivel en lo que respecta a las ciencias naturales. Además es el órgano consultivo supremo en cuestiones de asesoramiento científico y tecnológico y tiene a su servicio la infraestructura nacional más importante y el centro de desarrollo en materia de ciencias naturales y de alta tecnología más importante del país.⁸⁰

El Comité de Biodiversidad de CAS supervisa las operaciones de biodiversidad de las e-infraestructuras en China. El Banco de germoplasma de especies salvajes es una de las once e-infraestructuras de investigación que gestiona la Academia de Ciencias china. Adicionalmente, en CAS se encuentra el Centro Mundial de Microorganismos. Genera una serie de estándares para que los institutos puedan colaborar, estar interconectados.

⁸⁰ Para más tener más información, puede consultarse <http://english.cas.cn>. (Consultado por última vez el 1 de abril de 2017).

Capítulo 5. Coordinación para poder realizar análisis globales en la escala espacial y temporal

El CDB quiso ser un paraguas en lo que respecta a la biodiversidad, como lo fue, en su momento, el IPCC, que por diversas cuestiones (la politización del SBSTTA, la falta de consenso, la incapacidad de generar indicadores para medir la pérdida de biodiversidad, etc.) fracasó en su objetivo que no era otro que frenar la pérdida de biodiversidad. A pesar de no haber logrado el objetivo que se había propuesto, fue un convenio de referencia en lo que a biodiversidad se refiere, al haber logrado un cierto consenso, cuando menos, en cuanto al punto de partida.

Tras el fracaso del CDB empezó a fraguarse el IPBES, una nueva herramienta al servicio del CDB. A pesar de ello, prácticamente todos los países estaban representados en él. Fue constituido como un panel internacional, similar al de cambio climático, creado porque había resultado insuficiente el CDB ya que le faltaba información. A diferencia del GEO BON donde solo están los países del G20, todos los países del mundo forman parte del IPBES.

Se puso de relieve la existencia de cuatro áreas bien diferenciadas: la de políticas, la de investigación, la de asesoramiento y la de observaciones. En cada una de ellas existían distintos instrumentos, como se verá a continuación (Larigauderie, 2015). El CDB quedó enmarcado en las políticas; el IPBES, de alguna manera, representaba la parte de asesoramiento científico que la politización del SBSTTA no permitió que existiera en el CDB. El GEO y sus derivados quedaron como los responsables de las observaciones de la Tierra tanto en las nueve áreas del GEOSS como en el GEO BON, más centrado en los temas de biodiversidad. Finalmente, el GBIF trabajaba con los datos para ponerlos al alcance de todos los científicos (Larigauderie, 2015).

Por consiguiente, nacieron como organizaciones distintas y no todas hacían frente a las mismas necesidades (Larigauderie, 2016). De hecho, puede decirse que son complementarias y generan sinergias habida cuenta de que se mueven en marcos diferentes.

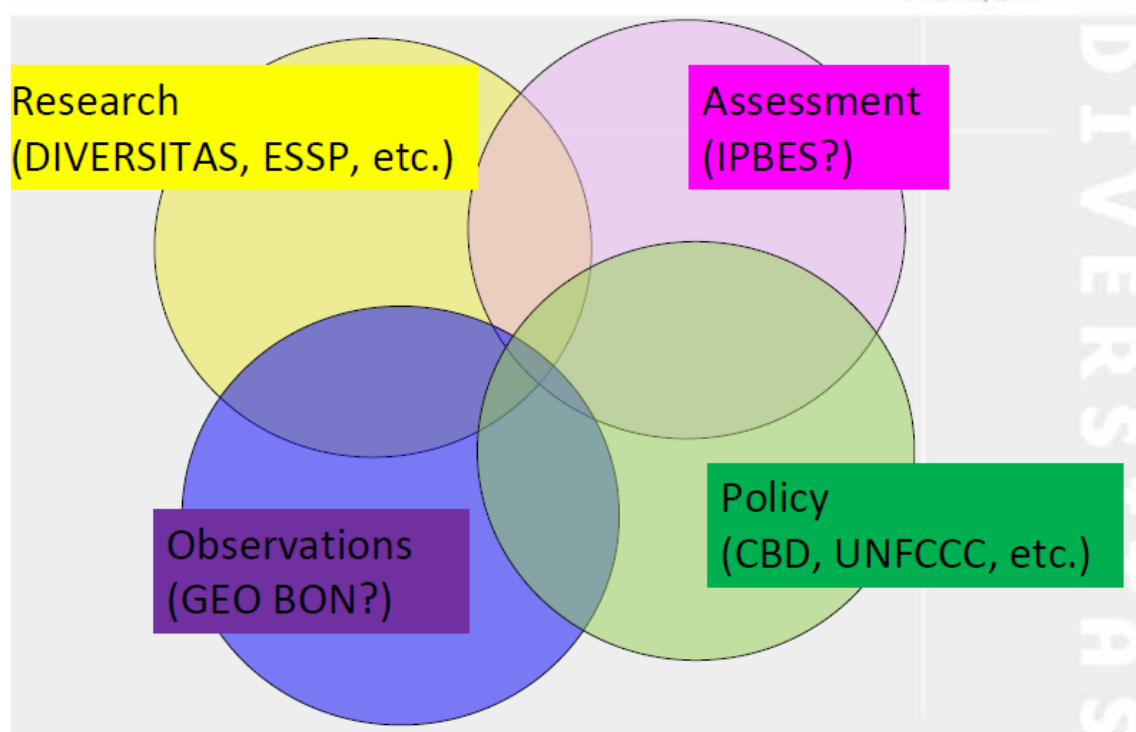


Figura 92. Como interactúan los distintos instrumentos. Fuente: Anne Larigauderie⁸¹.

El hecho de que cada una naciera respondiendo a una necesidad concreta permite la existencia de superposiciones de tareas. Es muy posible que, en el futuro, las herramientas consigan irse integrando construyendo una superestructura que coordine todo para optimizar las interacciones. Por el momento, es importante ir optimizando las interacciones entre ellas.

Al fin y al cabo, la gestión de la gobernanza de la biodiversidad depende de la existencia de información fiable (Dietz *et al.*, 2003). Sin lugar a dudas, la tendencia es la globalización de la gobernanza, con la finalidad de poder hacer frente a los nuevos problemas de índole global (Bolton, 2001). Esta afirmación es hoy aún más cierta que cuando Bolton la pronunció y la tendencia es ir hacia allá como demuestra el cuadro de Larigauderie.

Sin embargo, la globalización del análisis no debe separarse de la perspectiva de las informaciones regionales. Sigue siendo necesario, por un lado, recabar los datos a pequeña escala y, por otro, permitir el acceso a los mismos de forma global. De hecho, uno de los objetivos clave de los BON era movilizar los datos dando cobertura a los vacíos existentes. Los BON pueden contribuir de forma eficiente a la movilización de datos y al procesamiento de los que ya existan haciendo que desaparezcan las barreras y que los datos puedan encontrarse y sean accesibles y

⁸¹ Recuperado de: <http://www.unesco.org/mab/doc/iyb/scConf/Larigauderie.pdf>.

digestibles. Por consiguiente, los BON son un pilar clave para construir interfaces científico-políticas en el ámbito medioambiental y su papel se extiende a traducir los requerimientos políticos a investigación y seguimiento, catalizando de esta manera intercambios bidireccionales entre ciencia y política (Wetzel *et al.*, 2015).

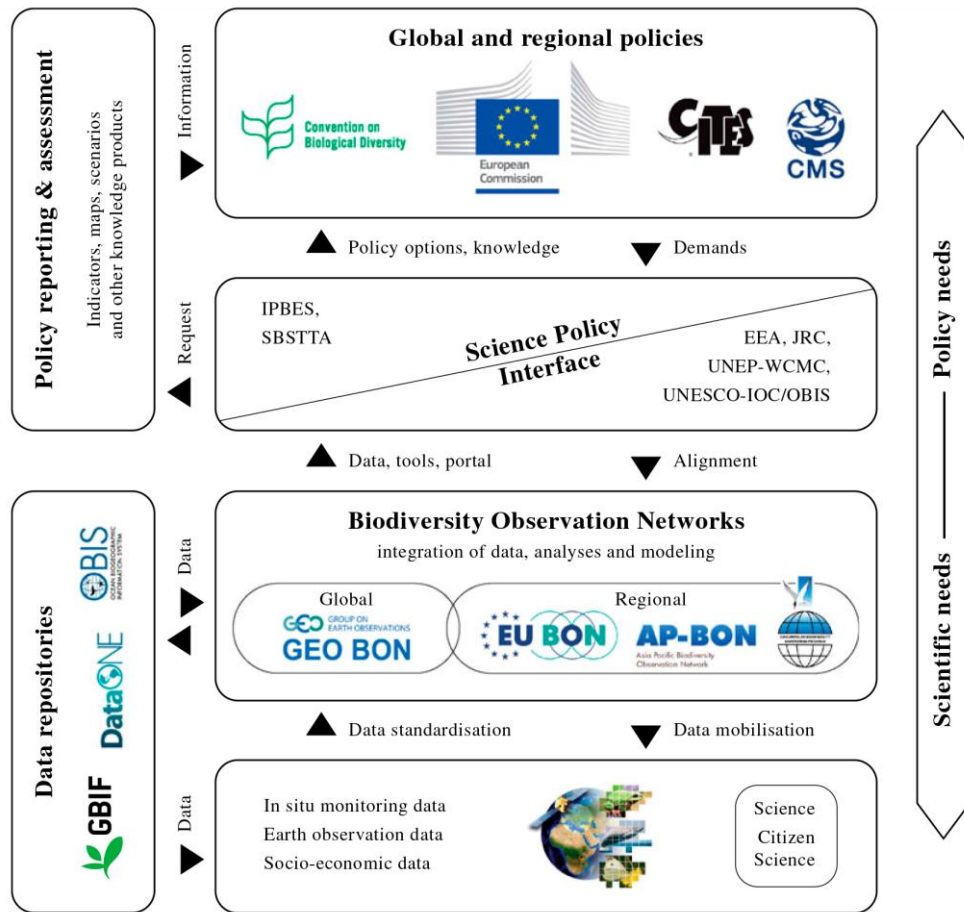


Figura 93. Ejemplos de movilización de datos, en este caso del EU BON. Refleja muy bien la interrelación entre las diferentes herramientas existentes. Fuente: Wetzel *et al.*, 2015.

Sin lugar a dudas, una de las tareas fundamentales de los BON era lograr la interoperabilidad dentro de los conjuntos de datos. Como puede verse en el gráfico, las distintas iniciativas han ido dando coherencia a todo el proceso y, sin perder de vista la necesidad de depurar, algunas, de hecho, muchas de ellas, están interrelacionadas y son interdependientes.

Es posible que, alguna, con el paso del tiempo, deje de ser necesaria o resulte absorbida por otra. A día de hoy, los datos de biodiversidad todavía están muy dispersos y separados, por lo que todas estas iniciativas son necesarias para contribuir a la estructuración de los mismos y a mejorar su accesibilidad.

Algunos autores se refieren a la gobernanza de la biodiversidad como una torre de Babel, tal vez como consecuencia de su fragmentación, señalan Soberón y Townsend (2015). El IPBES, prosiguen los mismos autores, se creó como una

forma de gestión del elemento de evaluación del espectro de políticas científicas elaborado para el CDB y otros acuerdos multilaterales para el medio ambiente.

Es casi un axioma en gobernanza de la biodiversidad que sus características sufren cambios en función del nivel donde se ponga el énfasis. Es más, las características de la biodiversidad son diferentes en función de la escala de la que se esté hablando, puesto que la escala es la que determina la identidad de los *stakeholders* y, por ende, sus preocupaciones y centro de interés. De alguna manera, este análisis justifica la coexistencia de órganos que actúen a distintas escalas. La clave es la interdependencia que lleva a la necesidad de construir *roadmaps* para posibilitar el avance.

1. LA CONSTRUCCIÓN DE ROADMAPS PARA FAVORECER EL TRABAJO CONJUNTO:

LOS PROGRAMAS DE PLANIFICACIÓN Y MODELIZACIÓN

Es necesario realizar un trabajo conjunto para facilitar que todo funcione, es decir, para que, de alguna manera, se alineen los programas de planificación y modelización. Sin embargo, todavía no se ha llegado a ese punto. El análisis de dos proyectos diferentes realizados en el marco de las subvenciones de la Unión Europea va a ser una forma de catalizar de todo lo anterior a la aplicación práctica en una situación real.

1.1. Una aplicación práctica de todo lo anterior a un problema y cómo se ha trabajado en el proyecto CREaTIVE-B

CREaTIVE-B son las siglas de Coordination of Research e-Infrastructures Activities Toward an International Virtual Environment for Biodiversity, un proyecto auspiciado por el 7.º Programa Marco en el que participan Universeit de Amsterdam (UvA), Cardiff University (CU), Maat France (Maat), Consiglio Nazionale delle Ricerche (CNR), Universidad de Alcalá (UAH), Comunita Ambiente (CA) y Centre National de la Recherche Scientifique (CNRS) y que arrancó en 2011.

Se invitó a participar a e-infraestructuras de varios lugares del mundo como *liaison partners*: Atlas of Living Australia (ALA), DataONE (Estados Unidos), NEON (Estados Unidos), Centro de Referencia de Informação Ambiental (CRIA) (Brasil), South African National Biodiversity Institute (SANBI), la Academia de Ciencias de China (CAS), el GBIF (global), World Federation of Culture Collections (WFCC), el GEO BON y, por supuesto, a la europea, LifeWatch, cuya creación, como ya se vio, arrancó también en 2011, aunque la decisión de crear el ERIC no se ha adoptado hasta este año 2017, fecha en la que nace realmente la e-infraestructura.

Por lo tanto, se trataba de e-infraestructuras sujetas a distintas regulaciones nacionales (Brasil, Sudáfrica, Estados Unidos, China, Australia), internacionales (el GEO BON) y globales (el GBIF). El objetivo de este proyecto era contribuir a la coordinación de las actividades de las diferentes e-infraestructuras de biodiversidad generando un espacio virtual internacional donde pudieran interactuar.

Por consiguiente, CREaTIVE-B es un espacio de colaboración entre infraestructuras que permitirá al LifeWatch ERIC interactuar con distintas e-infraestructuras de investigación en biodiversidad y ecosistemas en otras partes del mundo (Grant Agreement, 288441, 4). La finalidad más inmediata era la definición de un *roadmap* que analizara la interoperabilidad en tres niveles:

1. La *participación de la comunidad*, con la finalidad de incluir y atender las peticiones y necesidades de las comunidades científicas usuarias de las e-infraestructuras.
2. La *tecnología*, en lo que respecta a los datos, nuevas tecnologías y servicios de e-ciencia.
3. La *legal* y gobernanza, en relación con la propiedad y los derechos de acceso a los datos y las políticas globales de coordinación.

Por lo tanto, el proyecto pretendía catalizar una colaboración a nivel mundial dentro del área de la biodiversidad con la finalidad de encontrar la forma de coordinar estas e-infraestructuras para que pudieran interactuar. Más allá de este primer objetivo, estaba el de contribuir a los objetivos del GEOSS.

Este proyecto también tiene como finalidad contribuir a las prioridades que el G8 fija en la *Carta di Siracusa* (G8, 2009) con la finalidad de reforzar la cooperación para alcanzar una mayor supervisión de la biodiversidad, y conseguir así datos fiables, comparables e interoperables y desarrollar planteamientos globales para intercambiar conocimiento científico, mejores prácticas, tecnologías e innovación según las organizaciones existentes y los mecanismos que promueven investigación integral y enfocada y la capacidad de construir una evaluación sobre diversidad biológica, servicios de los ecosistemas y evaluación ambiental mundial.

Uno de los puntos de arranque del proyecto es el análisis de la interoperabilidad en dos niveles: el de la interoperabilidad técnica y el de la interoperabilidad legal. La interoperabilidad técnica se estudió en el *Deliverable 3* (Hardisty y Manset, 2014). El problema de base radicaba en evaluar si con los estándares técnicos utilizados era posible crear puentes (*interfaces*) entre las distintas bases de datos que permitieran acceder a los datos capturados en cada una de ellas. Si bien en un principio podía parecer que había limitaciones de interoperabilidad técnica, el informe concluyó que

no existían especiales trabas para facilitar la accesibilidad de los datos y el trabajo conjunto de todas las e-infraestructuras desde el punto de vista técnico.

De este análisis preliminar resultó que las nueve e-infraestructuras parecían tener un buen potencial de interoperabilidad teniendo en cuenta la manera en la que daban acceso a los datos de biodiversidad, aplicaciones y recursos relacionados (Hardisty y Manset, 2014). Parecía que perseguían objetivos similares en términos de modelos de negocio, involucración en políticas e industria.

Más allá de esto, eran e-infraestructuras complementarias desde el punto de vista geográfico y de los temas que contemplaban. No obstante, sí que se planteaban algunas cuestiones en relación con la integración de datos en el sentido de dar consistencia a los términos para poder relacionarlos. Otra cuestión que puso de relieve es el concepto de la sostenibilidad en el tiempo de estas e-infraestructuras, dado que muchas de ellas se sostienen ahora con fondos públicos.

La interoperabilidad legal se estudió en el *deliverable* 4 (Alonso García, 2014), analizándose todos los documentos legales que sirvieron de punto de arranque de las distintas e-infraestructuras con la finalidad de ver si los estándares adoptados tenían que ver con requerimientos legales o existían restricciones a la hora de modificarlos. La principal preocupación era detectar las barreras que pudieran dificultar la interacción de las distintas bases de datos. Como se verá después en el punto de la interoperabilidad en el marco de CREaTIVE-B, las barreras no fueron de la entidad que se anticipó. Por lo tanto, resultó que no se daban barreras legales para que las e-infraestructuras pudieran trabajar de forma conjunta si existía la voluntad de colaboración.

Uno de los enfoques fundamentales del proyecto es explorar de qué manera la cooperación y la interoperabilidad a gran escala de las infraestructuras de investigación en todo el globo pueden soportar los retos de la investigación realizada en biodiversidad y ecosistemas.

Uno de los logros de este proyecto fue alcanzar el compromiso de las infraestructuras de investigación de continuar con esta colaboración una vez concluido el proyecto con la finalidad de seguir avanzando en el progreso científico a la hora de comprender y predecir la complejidad de los sistemas naturales. Después de haber trabajado conjuntamente en la implementación de las recomendaciones de este *roadmap* (Alonso García *et al.*, 2014) los datos y posibilidades muestran que las e-infraestructuras de investigación participantes están mejor preparadas para hacer frente a los retos en lo que respecta a la biodiversidad y a los ecosistemas de lo que se esperaba al principio.

1.2. El roadmap

Las recomendaciones del *roadmap* perseguían promover que los usuarios se comprometieran y aportaran valor centrándose en apoyar los objetivos globales y comunes de investigación, así como en desarrollar un conjunto de tecnologías avanzadas y en que se involucrasen los científicos en lo que afecta a las actividades que incluyan observación medioambiental y seguimiento. A pesar de que las infraestructuras de investigación tienen un nivel satisfactorio de potencial interoperabilidad, continúan existiendo barreras para una interoperabilidad global. Resulta necesario realizar acciones con la finalidad de ir promoviendo la mejor comprensión y el valor de la interoperabilidad de las e-infraestructuras de investigación, el desarrollo de mecanismos de coordinación que consigan la interoperabilidad y la búsqueda cada vez mayor de estándares.

El reto era la creación de un mercado científico que permitiera a los usuarios beneficiarse de los flujos de servicios tal como se usan en las e-infraestructuras de investigación involucradas en el proyecto. Compartir datos y herramientas dentro de estos flujos de origen variable en relación con los autores y propietarios⁸² exigía cuidados y eficientes configuraciones con la finalidad de que los usuarios pudieran beneficiarse de los recursos combinados sin tediosas restricciones. Esto es todavía más importante con el cada vez mayor procesamiento automático de datos realizado de máquina a máquina.

Las e-infraestructuras participantes acuerdan que cada una explorará las nuevas oportunidades de comenzar a aplicar las recomendaciones.

⁸² Cada vez más se está cuestionando el término de propiedad aplicado a los datos.

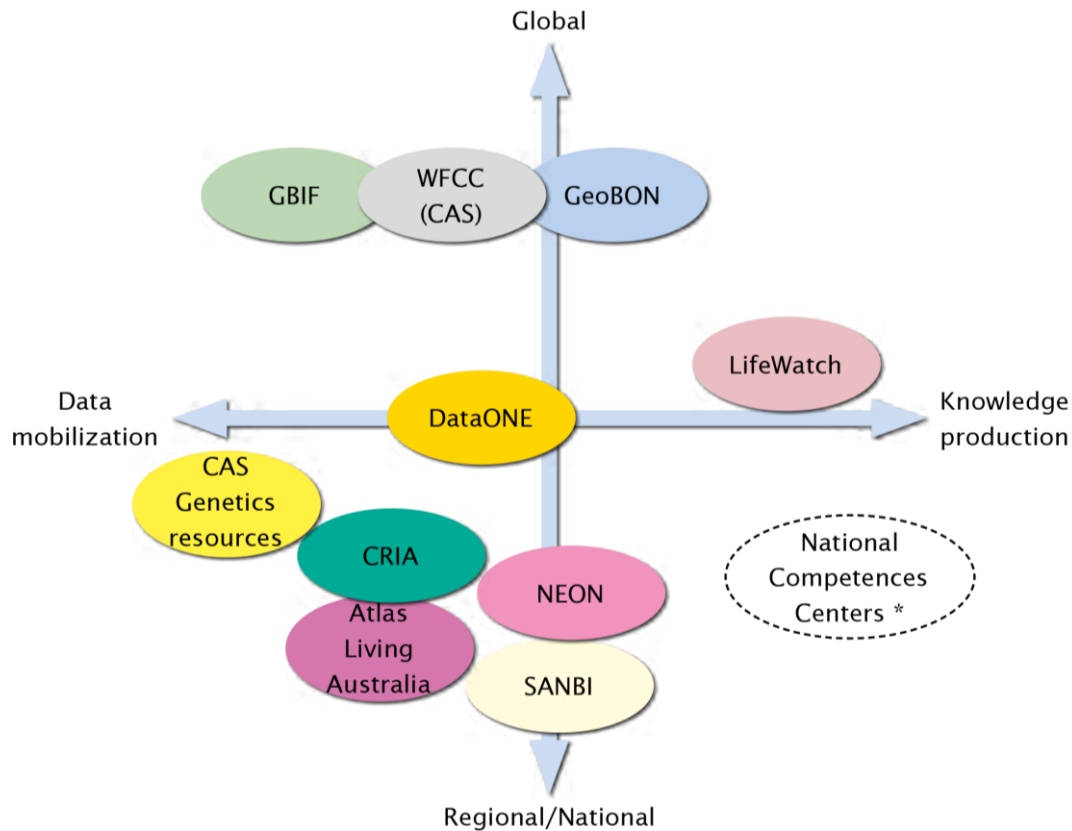


Figura 94. How we fit together: global view. Fuente: CREaTIVE-B.

Los ejes de este gráfico representan distintas dimensiones que caracterizan las distintas infraestructuras de investigación. El eje horizontal hace referencia a sus misiones operativas. En el lado izquierdo, están las distintas estructuras que tienen como misión movilizar los datos ofreciendo su almacenamiento, su intercambio y servicios de acceso a ellos. En el lado derecho están las pocas e-infraestructuras que apoyan el uso de datos para su análisis y modelización. Las e-infraestructuras que los movilizan ofrecen cada vez más estos servicios de procesamiento. El eje vertical muestra en la parte inferior las e-infraestructuras fundamentalmente operacionales en el ámbito nacional y regional; arriba se encuentran las que dan servicios exclusivamente globales.

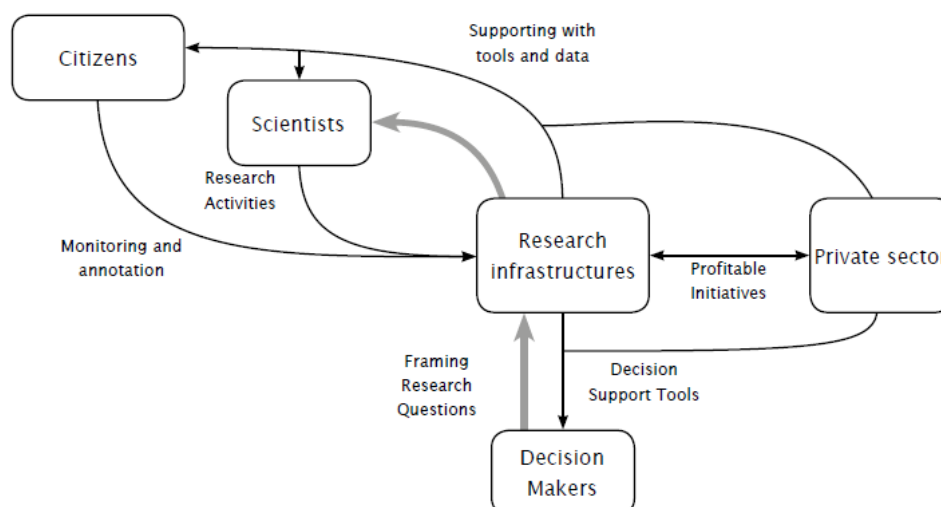


Figura 95. Rol de las e-infraestructuras en la facilitación de la colaboración. Fuente: Alonso García *et al.*, 2014.

En el marco de CREaTIVE-B, se han identificado una serie de retos a los que las e-infraestructuras que colaboran tienen que hacer frente. El primero de ellos es la necesidad que reduzcan las barreras con la finalidad de que los científicos usuarios puedan beneficiarse de las ventajas de infraestructuras «virtuales». Los entornos virtuales amigables deben simplificar los procesos de acceso, de modo que los investigadores les dediquen menos tiempo. Es necesario elaborar materiales y actividades formativas, pero no es esta la única solución.

El segundo de ellos es comprometer a la comunidad científica con la validación de datos para potenciar la calidad de los mismos. Este *feedback* es un complemento esencial para los mecanismos de validación automatizada dentro de todas las infraestructuras de investigación. La comunidad de investigación de biodiversidad necesita estar motivada y empoderada para poder desarrollar estas tareas en línea y de manera colaborativa. Este entorno de validación no existe en la actualidad, pero forma parte de la columna vertebral fundamental de la infraestructura de biodiversidad en el punto 20 de la *Visión Decenal de la Bioinformática*.

El tercer reto al que se enfrentan las e-infraestructuras consiste en generar fuertes redes de investigadores de biodiversidad y ecosistemas construyendo laboratorios virtuales que permitan la creación de grandes grupos de investigación para cooperar de forma remota en grandes retos. Esto debe terminar en herramientas comunes como parte de servicios a gran escala para comunidades estructuradas.

En el proyecto de CREaTIVE-B, se evaluaron los distintos niveles de interoperabilidad para identificar posibles barreras para que las e-infraestructuras pudieran interactuar. El punto de arranque fue analizar barreras técnicas y legales.

General overview

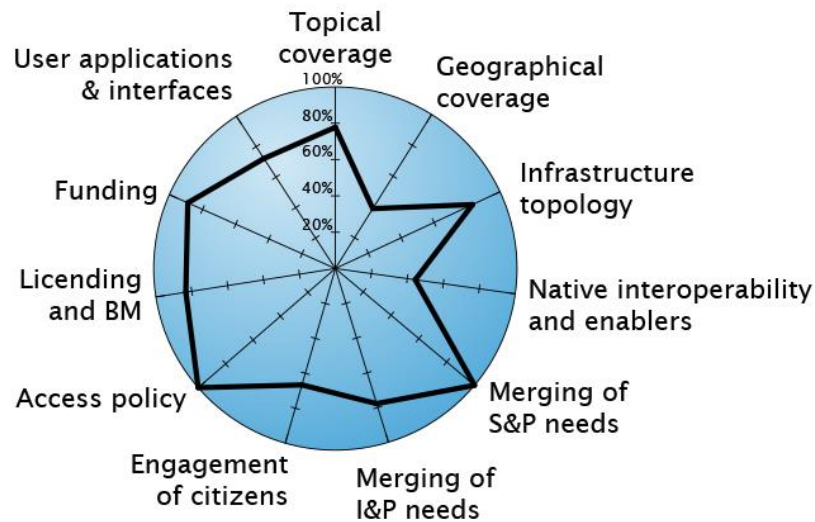


Figura 96. Evaluación del potencial de lograr la interoperabilidad. Visión general. Fuente: Alonso García *et al.*, 2014.

En la realización del proyecto se hace una evaluación de la potencialidad global de interoperabilidad. Como puede verse, las e-infraestructuras muestran un nivel satisfactorio de interoperabilidad, particularmente en lo que respecta al acceso a los datos de biodiversidad, a las aplicaciones disponibles y a los recursos relacionados con esto. En términos de modelo de negocio, de industria, de implicación política y de planes de sostenibilidad global, todas las e-infraestructuras persiguen objetivos similares. Estos objetivos pretenden la creación de un entorno virtual internacional para investigación de ecosistemas y biodiversidad. Las distintas e-infraestructuras tienen una cobertura geográfica complementaria.

Service logic

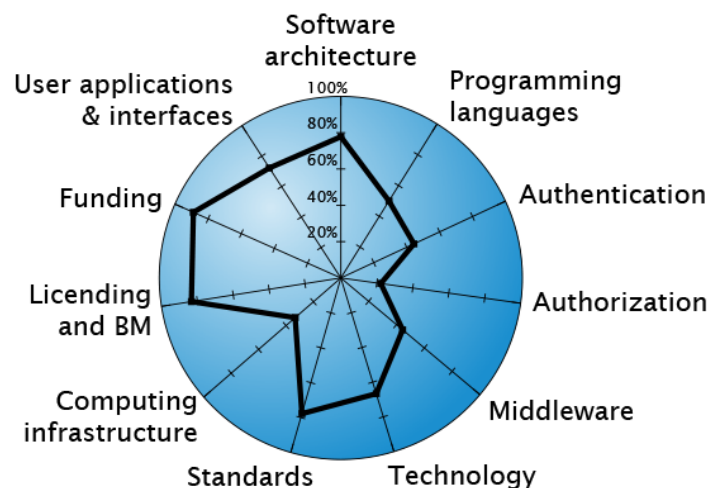


Figura 97. Evaluación del potencial de lograr la interoperabilidad. Lógica de los servicios.
Fuente: Alonso García *et al.*, 2014.

En este punto, en lo que respecta a la arquitectura, a los estándares, entre otros temas, es donde más diferencias existen entre las e-infraestructuras. Se han usado distintos programas intermedios, medidas de seguridad, lenguajes de programación, estándares, etc. Sin lugar a dudas, en la parte que más desarrollo necesita para que los sistemas sean sintáctica y semánticamente compatibles.

Data

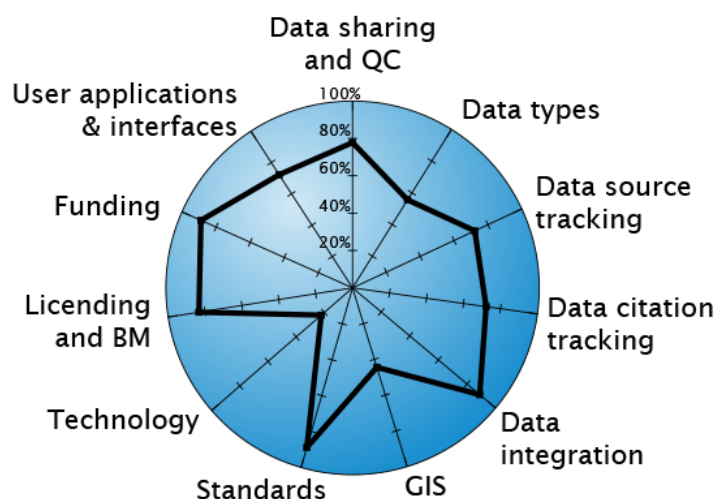


Figura 98. Evaluación del potencial de lograr la interoperabilidad. Datos. Fuente: Alonso García *et al.*, 2014.

Con este último gráfico queda demostrado que ya empezaban a existir estándares que comenzaban a enfocar las necesidades de integración de datos y de organización. Existían procesos de control de calidad y de intercambio de datos para lidiar con las colecciones de datos, y la trazabilidad se había convertido en una preocupación para las citas de los científicos y para hacer el seguimiento de los datos brutos de las investigaciones.

Por lo tanto, superar las barreras de la interoperabilidad global en las e-infraestructuras significaba:

1. Promover que se entienda el valor que tiene que las e-infraestructuras sean interoperables.
2. Hacer uso de la coordinación para lograr la interoperabilidad.
3. Enfatizar e incrementar la importancia de los estándares.
4. Resolver retos tecnológicos específicos.

CREaTIVE-B hizo hincapié, de esta forma, en el *roadmap* (Alonso García *et al.*, 2014) en el valor de hacer que las e-infraestructuras fueran interoperables. Para ello, era importante usar casos de estudio que ilustrasen la necesidad de interoperabilidad. En relación con la coordinación para lograrla, era necesario enfatizar la importancia de estándares, pues su inestabilidad era una barrera para la interoperabilidad. A efectos, de crear y consolidar estándares en este ámbito que contribuyesen a la estabilidad e interoperabilidad, la recomendación de CREaTIVE-B fue buscar casos de éxito en los que se hubiesen establecido estándares que permitiese publicar y promover uno relacionados con la biodiversidad.

En relación con la última de las cuestiones planteadas, la interoperabilidad técnica, se hicieron toda una serie de recomendaciones para que la comunidad de biodiversidad pudiese alinear y conectar sus servicios:

1. Desarrollar e-infraestructuras posibilitadoras, globales y federadas AAA, es decir, autenticadas, autorizadas y en las que se pueda confiar en los siguientes tres años.
2. Promover el uso de controles de calidad, de controles semánticos. Es necesario supervisar la calidad de los datos y metadatos para que puedan convertirse en información útil para los usuarios.
3. Promover el desarrollo, intercambio y uso de los *workflows* de servicios.
4. Crear un mercado científico para los servicios de biodiversidad.
5. Gestionar la procedencia de los recursos de las e-infraestructuras de investigación.

Por supuesto, el *roadmap* (Alonso García *et al.*, 2014) profundizó en todo lo que estos puntos requerían, los pasos necesarios y una estimación de tiempos de consecución. Este es uno de los proyectos, y ahora procede analizar el siguiente ejemplo de aplicación. En las reuniones de CREaTIVE-B empieza a abordarse ya la necesidad de que existan de las EBV ligadas al proyecto.

1.2.1. Un paso más, GLOBIS-B

GLOBIS-B (GLOBal Infrastructures for Supporting Biodiversity Research) es un proyecto auspiciado en el marco de Horizonte 2020, que, de alguna forma, quiso dar un paso más dentro del análisis iniciado en el marco de CREaTIVE-B (GLOBIS-B, 2015). La finalidad del proyecto es contribuir a desarrollar medidas clave que sustentan los indicadores globales que es necesario estudiar, reportar y gestionar en relación con los cambios de biodiversidad (Pereira *et al.*, 2013). Más concretamente, se centra en los potenciales servicios de infraestructuras que sustentan la investigación dedicada a las mencionadas EBV o variables esenciales de biodiversidad (Konijn, 2016). Persigue, así, promover la cooperación de las

infraestructuras de investigación y de los científicos centrados en el campo de la biodiversidad, con la finalidad de adelantar la implementación y cálculo de las variables esenciales de la biodiversidad.

Una de las cuestiones clave en la monitorización global de esta es cómo puede articularse la cooperación multilateral de los recabadores de datos, sus proveedores, los sistemas de seguimiento y las infraestructuras de investigación de biodiversidad en un ámbito global para dar soporte a la implementación armónica de EBV, que hasta este momento ha resultado complicado testar. Las fronteras de investigación en esta área exigen disponibilidad y accesibilidad de conjuntos de datos sustanciales con cobertura espaciotemporal suficiente.

La finalidad de GLOBIS-B es dilucidar de qué manera las infraestructuras de investigación que cooperan pueden contribuir a dicho objetivo centrándose en ofrecer los datos y sus flujos así como servicios computacionales para calcular las EBV (Konijn, 2016).

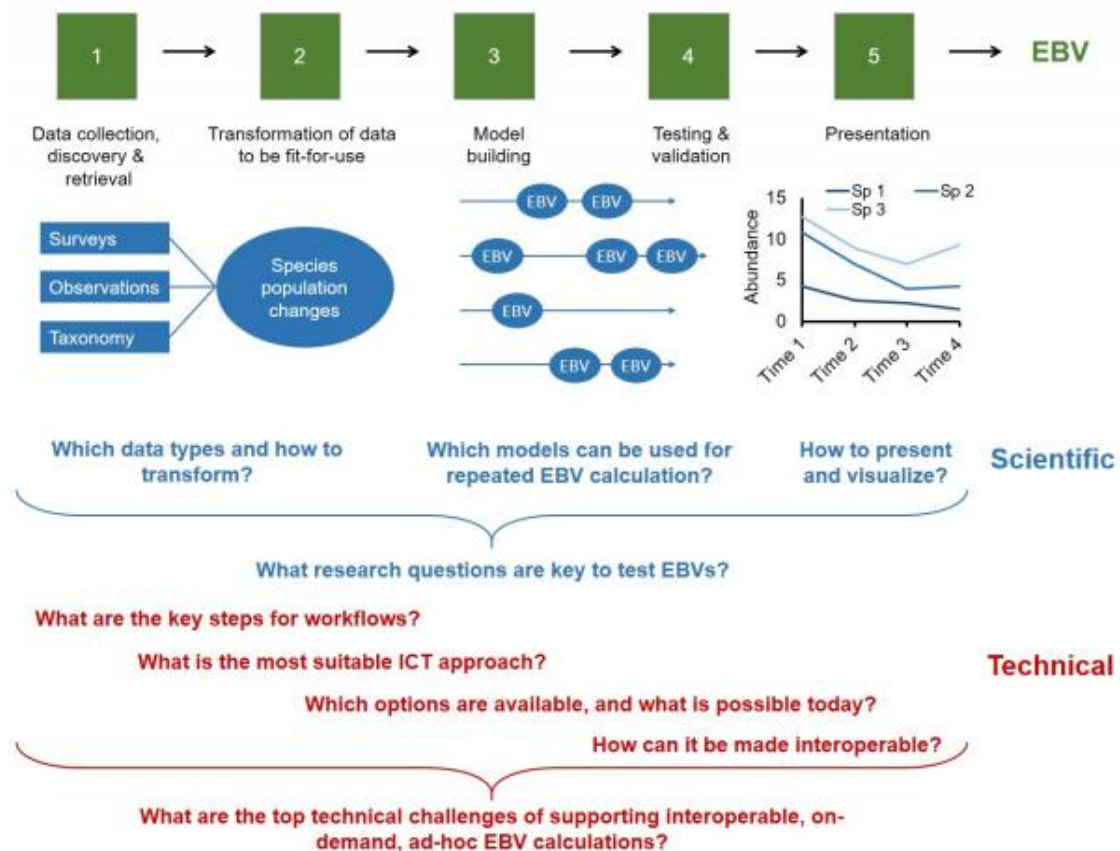


Figura 99. Posibles pasos para el cálculo de las Variables Esenciales de Biodiversidad (EBV, verde) y relacionadas preguntas y retos científicos (azul) y técnicos (rojo). Fuente: Konijn, 2016.

La duración prevista del proyecto es de treinta y seis meses con fecha de término en mayo de 2018. El objetivo es identificar los datos primarios requeridos y las herramientas de análisis, entre las que se encuentran identificar metodologías y

cuellos de botella jurídicos y técnicos para desarrollar una agenda de desarrollo de la investigación y de las e-infraestructuras para programar las EBV.

Por ello, pretende poner a trabajar conjuntamente a científicos de biodiversidad, operadores de e-infraestructuras globales y expertos de interoperabilidad legal con la finalidad de identificar las necesidades de investigación y de servicios de e-infraestructuras sobre las que se apoyan las EBV.

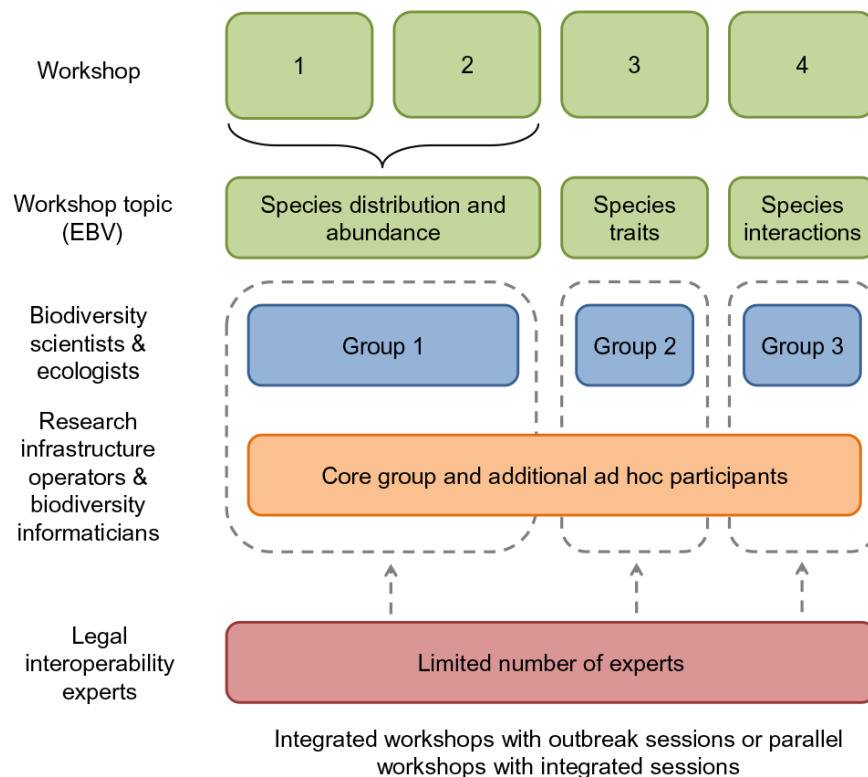


Figura 100. Perspectiva de los 4 talleres de GLOBIS-B relacionados con distribución de especies y abundancia (1 y 2), características de especies (3) e interacciones de especies (4). Fuente: Konijn, 2016.

La finalidad de los talleres es discutir y desarrollar un marco para implementar las EBV entre las infraestructuras a nivel mundial. Podrá hacerse mediante debates entre distintos grupos participantes, incluidos científicos de biodiversidad y ecologistas, informáticos de biodiversidad, tecnólogos y operadores de infraestructuras y expertos en interoperabilidad legal (Kissling *et al.*, 2015).

La finalidad de los dos primeros talleres fue identificar las opciones y los problemas asociados con el reto general: provisionar las infraestructuras de investigación para ofrecer funciones que den soporte a la generación y cálculo de las EBV. El primero de los talleres se organizó con sesiones integradas y en paralelo de ecologistas y científicos de biodiversidad (especializados en distribución, abundancia y seguimiento de especies), informáticos expertos en materia de biodiversidad y

operadores de infraestructuras. El segundo de los talleres consistirá en un seguimiento, algunos meses después y con esos mismos expertos, con la finalidad de proporcionar una actualización de las iniciativas que estén en marcha y garantizar que se escriba un estudio revisado por pares y que se publiquen los resultados clave como seguimiento al artículo de Pereira *et al.* (2013). El primero de los talleres ya se ha celebrado (Konijn, 2016). Para conseguirlo se ha organizado de la siguiente manera:

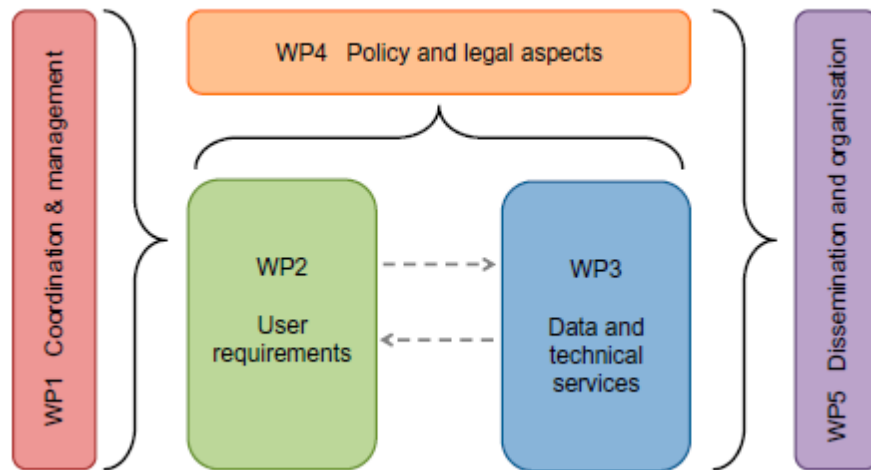


Figura 101. Presentación gráfica de los elementos del plan de trabajo de GLOBIS-B integrado por 5 paquetes de trabajo. Los paquetes de trabajo WP2-4 se encargan del contenido de los *workshops* mientras que los paquetes WP1 y WP5 son, fundamentalmente de carácter administrativo

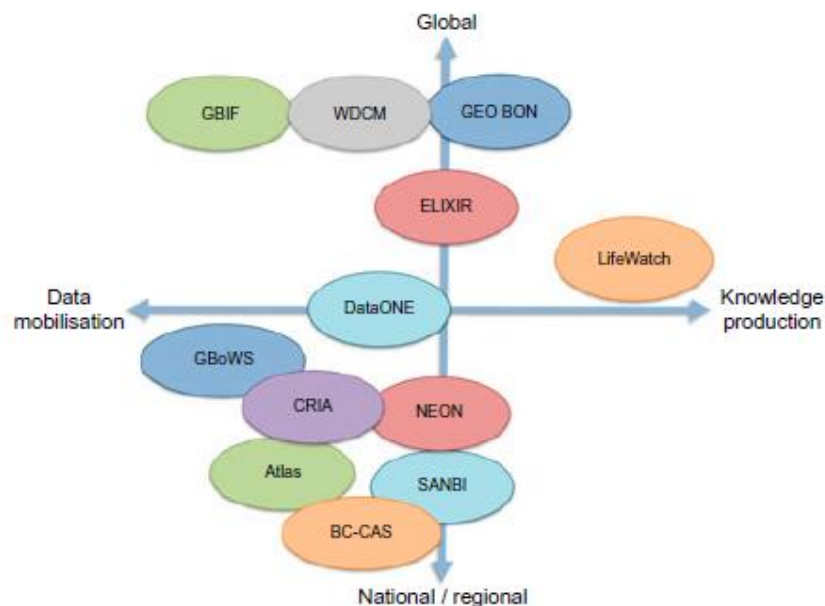


Figura 102. Caracterización de las infraestructuras de investigación que cooperan en el proyecto GLOBIS-B. Las infraestructuras de investigación son distintas en su ámbito geográfico (eje-y) y su foco en movilización de datos vs. producción de conocimiento eje-x. Esta imagen se ha hecho tomando como base otra realizada para el proyecto CREaTIVE-B. Fuente: Kissling *et al.* 2015.

El segundo de los talleres, celebrado en junio de 2016, tuvo como foco, en primer lugar, realizar un seguimiento del primero pero también trabajar la EBV de Poblaciones de Especies (es decir, en el estudio de las distribuciones de especies). El tercero de los talleres, celebrado en marzo de 2017, estuvo enfocado en las características de las especies. El último, celebrado en febrero de 2018, tuvo como tema central las interacciones entre las especies.

El objetivo final de GLOBIS-B era, por lo tanto, conseguir una infraestructura de apoyo global y mejorada para calcular las EBV que permita asesorar en el progreso hacia la consecución de los Aichi Targets de 2020 (Kissling *et al.*, 2015).

La finalidad de este proyecto era llegar a tres audiencias concretas y contribuir de este modo a la agenda de biodiversidad (Kissling *et al.*, 2015):

1. La comunidad de infraestructuras de investigación.
2. La comunidad científica de biodiversidad.
3. Los organismos internacionales principales en lo que respecta a políticas de biodiversidad (el IPBES, el CDB).

Al fin y al cabo, la finalidad del seguimiento de la biodiversidad global es medir las respuestas a los cambios medioambientales (Proença *et al.*, 2016).

Después de todo este análisis, se puede decir que el nuevo marco de la sociedad del conocimiento, de la revolución digital, ha supuesto un cambio de paradigma en los temas de biodiversidad. Los distintos instrumentos desarrollados para frenar la pérdida de biodiversidad no han tenido el impacto esperado. Sin embargo, más allá de eso, se identifica la necesidad de empezar a medir de forma más correcta y exacta la pérdida de biodiversidad a la que se está haciendo frente.

Parece que el camino empezado puede llegar a un buen fin, aunque la solución esté todavía en proceso de elaboración. Toda la problemática derivada de las EBV merece análisis más profundo.

1.3. Las variables esenciales de biodiversidad o EBV

GLOBIS-B da servicio al GEO BON, que introduce el concepto de EBV como una de las áreas de beneficios para la sociedad del GEOSS (Pereira *et al.*, 2013). Las EBV son medidas necesarias para estudiar, reportar y gestionar el cambio de la biodiversidad.

Se produce un proceso de transición de los Aichi Targets contruidos en el marco del CDB y que, como ya se vio, no tuvieron el impacto esperado, a las EBV. De hecho, un posible análisis puede hacerse a través de la llamada observación de la tierra. Es importante que la comunidad científica contribuya en el ámbito de la generación de las EBV a través de los distintos foros. Por ello, la observación se convierte en una herramienta para desarrollar las EBVs (O'Connor *et al.* 2015).

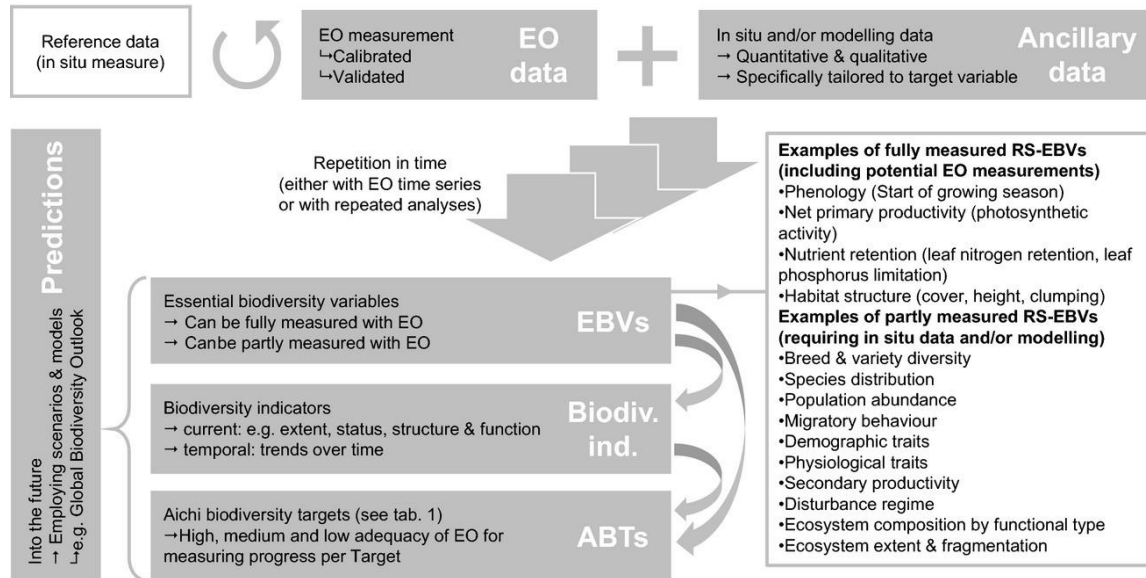


Figura 103. Camino propuesto para pasar de la observación de la tierra a las EBVs, a los indicadores de biodiversidad y a los Aichi Targets con ejemplos de observación de la tierra para hacer mediciones parciales o completas de las EBVs. Fuente: O'Connor *et al.* 2015.

Se considera que las EBV proporcionan un enfoque holístico y trazable para monitorizar el estado de la biodiversidad (Pereira *et al.*, 2013). Pueden ser cruciales para realizar una estimación y una derivación robusta de indicadores de biodiversidad con la finalidad de evaluar el progreso hacia la consecución de los Aichi Targets. Para el GEO BON, las EBV tienen que reunir una serie de criterios:

- Tener la capacidad de capturar escalas y dimensiones de la biodiversidad críticas.
- Ser biológicas.
- Tener una variable de estado (en general).
- Ser sensibles al cambio.
- Ser agnóstico del ecosistema (siempre que sea posible).
- Ser técnicamente factibles, viable desde el punto de vista económico y sostenible en el tiempo.

Tiene que ser posible medir y construir estas EBV globalmente. Sería ideal poder obtenerlas de la integración de las teleobservaciones con las observaciones *in situ*.

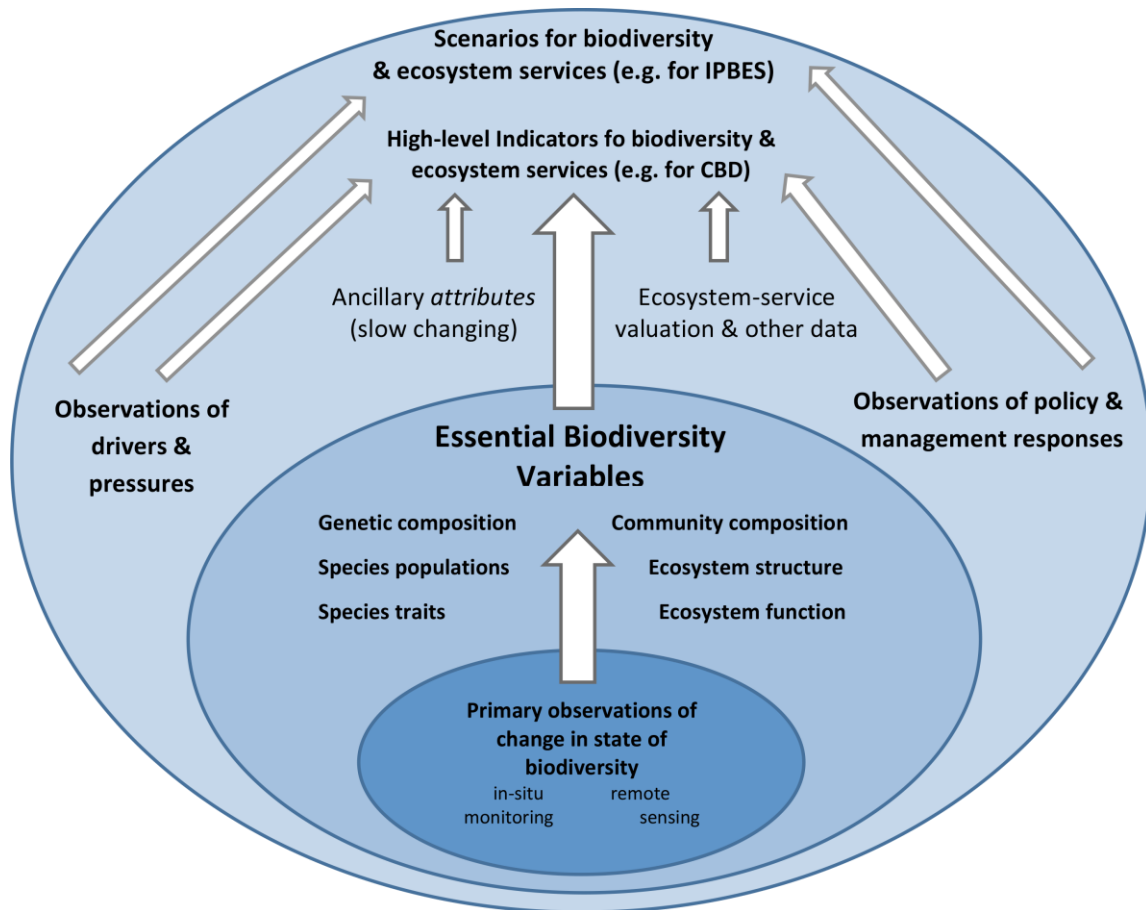


Figura 104. GEO BON. Fuente GEO BON. Fuente: GEO BON.⁸³

Para el GEO BON, las EBV, que se encuentran entre la observación primaria y los indicadores, tienen que convertirse en la ventana de los sistemas de observación de la biodiversidad para los gestores, los investigadores y los que tienen que tomar las decisiones a distintos niveles (GEO BON Management Committee, 2017).

El GEO BON (Pereira *et al.*, 2013) ha propuesto un total de veintidós EBV agrupadas en seis clases: composición genética, poblaciones de especies, características de especies, composición de comunidades, funcionamiento de ecosistemas y estructura de ecosistemas.⁸⁴

⁸³ Recuperado de: <http://geobon.org/essential-biodiversity-variables/what-are-ebvs>.

⁸⁴ Disponible en: <http://GEOBON.org/essential-biodiversity-variables/ebv-classes-2>. (Consultado por última vez el 3 de agosto de 2016).

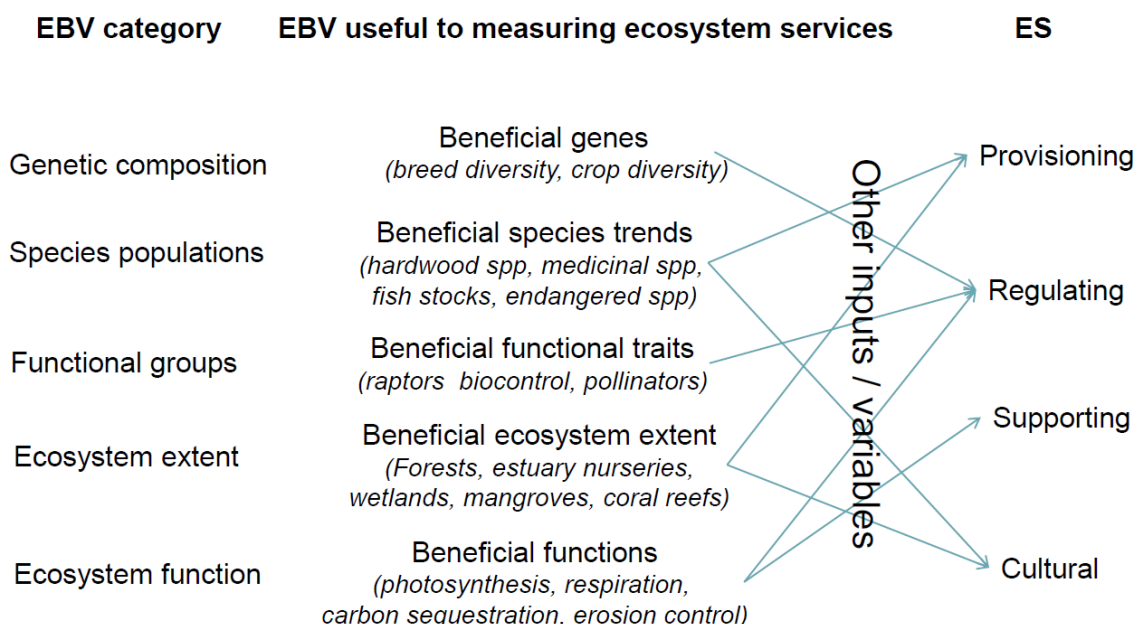


Figura 105. Variables Esenciales de Biodiversidad. Fuente: Pereira, 2013.

GLOBIS-B se centra en los datos relacionados con distribución de especies y volúmenes (clase EBV «poblaciones de especies»), características ecológicas (clase EBV «características de especies») e interacciones de especies (clase EBV «composición de comunidades»).

Es necesario un sistema global y armonizado de observaciones con la finalidad de tener informados a los científicos y a los que elaboran las políticas. Pereira *et al.* (2013) observan la complejidad del cambio de la biodiversidad que se resumen en este gráfico.

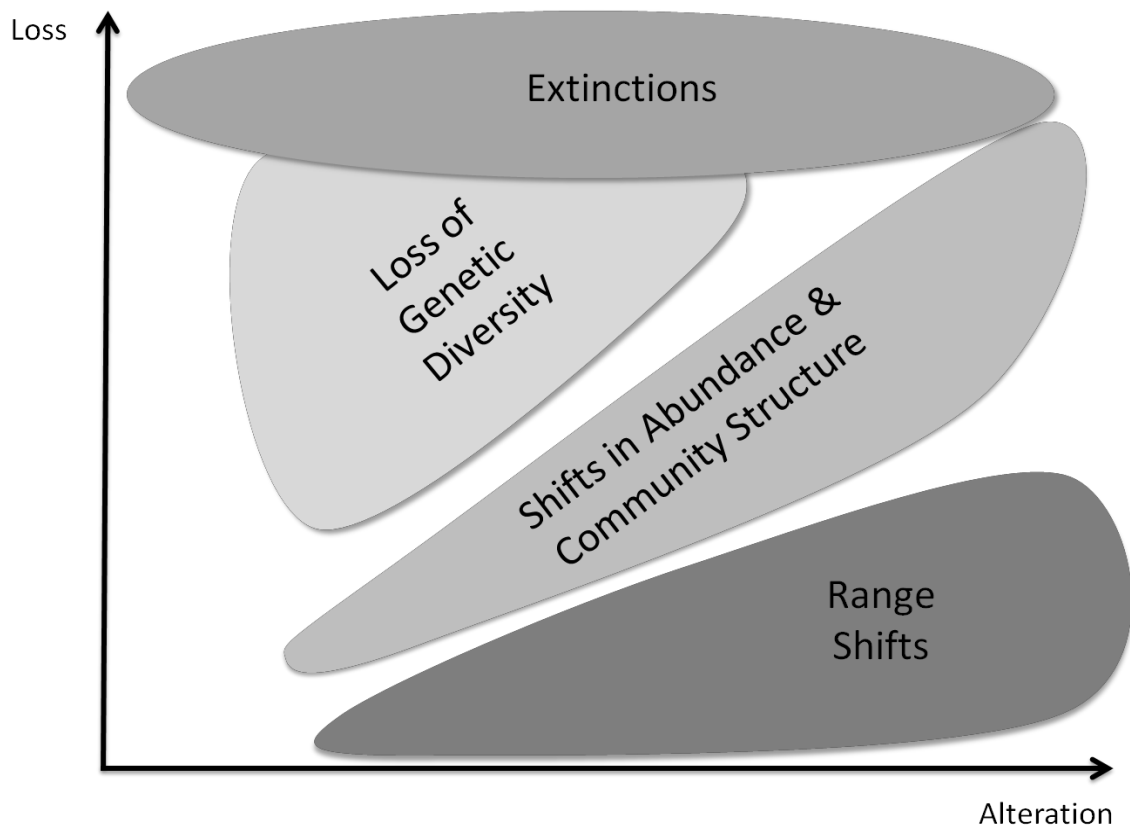


Figura 106. Diagrama conceptual que ilustra la intensidad de la pérdida y alteración asociadas con las distintas dimensiones del cambio de la biodiversidad: extinciones, pérdida de diversidad genética, cambios en la abundancia y estructura de la comunidad y zonas de ajuste. Fuente: Pereira *et al.*, 2012.

Los autores ponen de relieve las lagunas que existen en lo que respecta a la monitorización de la biodiversidad haciendo hincapié en que no existe ningún sistema de observación global armonizada que permita informar de forma puntual de los datos de cambio en la biodiversidad.

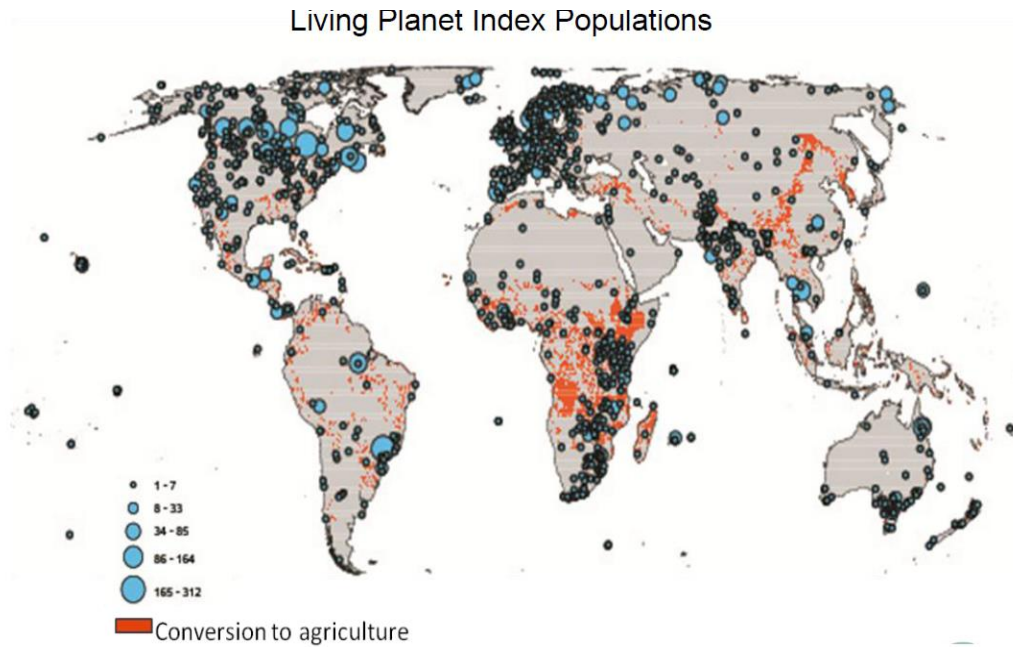


Figura 107. Índice de poblaciones de vida en el planeta. Fuente: Pereira, 2013.

Más allá de la inexistencia de un sistema global, Perera *et al.* (2013) ponen de manifiesto la inexistencia de consenso sobre las variables que es necesario monitorizar y compartir. Una vez más, los sistemas de biodiversidad se van a las estructuras de cambio climático que tratan de reproducir.

La finalidad de las EBV es ayudar a las comunidades de observación a armonizar la monitorización identificando las variables que es necesario medir, de las que se han propuesto más de cien. Sin embargo, la finalidad de estas variables tiene que ser capturar los datos de las medidas esenciales que permitan evaluar las mayores dimensiones del cambio de biodiversidad.

Por lo tanto, Pereira *et al.* definen la EBV como una medida necesaria para estudiar, reportar y gestionar el cambio en biodiversidad. Encajan potencialmente en esta definición cientos de variables. Por esta razón, han realizado un trabajo de análisis con la finalidad de identificar las que cumplen los criterios de escalabilidad, sensibilidad temporal, viabilidad y relevancia.

En lo que respecta a la coordinación de las e-infraestructuras, ha quedado suficientemente puesto de relieve en el GBIO la complejidad que existe para construir ese espacio virtual interconectado.

Existe una multiplicidad de retos para la consecución de la implementación global de las EBV. Su cálculo exige que concurren una serie de elementos (Kissling *et al.*, 2015), incluidos los siguientes:

1. La disponibilidad de los datos de biodiversidad en escalas espaciales, temporales y actuales relevantes.

2. Herramientas adecuadas de análisis y servicios para calcular, analizar, visualizar e interpretar EBV.
3. La interoperabilidad de las infraestructuras de investigación existentes para integrar los datos de biodiversidad y para, de forma continuada, calcular las EBV a demanda.

En primer lugar, los datos de biodiversidad para calcular las EBV tienen que estar disponibles y haberse descubierto, por ejemplo, en los ámbitos de la biodiversidad y de la observación de la Tierra y sus portales de datos (Pereira *et al.*, 2013).

En segundo lugar, se necesitan herramientas y servicios para el cálculo, análisis, visualización e interpretación de las EBV. Por lo tanto, como ponen de manifiesto Pereira *et al.* (2013), para que se puedan implementar EBV de manera uniforme en el mundo, es necesaria la cooperación global de los científicos de biodiversidad y los operadores de e-infraestructuras, para facilitar la interoperabilidad de los datos y las herramientas de procesamiento asociadas.

Por lo tanto, dos son los retos fundamentales a los que se enfrenta la implementación global de EBV: los de carácter científico y los de carácter técnico (Kissling *et al.*, 2015). Los de carácter científico están relacionados con el propio concepto de EBV, es decir: cómo se definen exactamente las EBV, qué datos de biodiversidad se necesitan, la disponibilidad de los mismos y de dónde se obtienen, de qué manera se puede acceder a los datos relevantes, cómo se pueden calcular las EBV, cuáles son los ámbitos espaciales, temporales y actuales relevantes y cómo de sensibles son las EBV a las variaciones que se produzcan en los datos que constituyen su base.

Estas son algunas de las cuestiones que apuntan Kissling *et al.* en su artículo (2015). Prosiguen diciendo que esto requiere identificar los datos que, de forma realista, se pueden recoger y de qué manera puede hacerse esto, cómo se puede mejorar la publicación y la accesibilidad de los datos y qué protocolos comunes (estandarizados) son necesarios para medir los datos y calcular las EBV.

El segundo conjunto de retos reside en la parte técnica. El desarrollo de las EBV requiere la cooperación global de las infraestructuras de investigación que generen *sets* de datos y variables comparables, lo cual lleva aparejada la interoperabilidad de las mismas (Kissling *et al.*, 2015).

Este reto de interoperabilidad tiene dos vertientes: una técnica y otra legal. La primera requiere la existencia de estándares y protocolos de IT ampliamente extendidos que permitan a las distintas e-infraestructuras interactuar. La finalidad no es otra que la de hacer posible la reutilización efectiva y fácil de los datos ecológicos.

La interoperabilidad legal —que se analizará con más profundidad en el apartado siguiente— requiere una autorización «máquina-máquina» más allá de los límites políticos e institucionales. Con frecuencia y como consecuencia de las variaciones en la procedencia de la autoría y propiedad de los datos, esto presenta dificultades y requiere la identificación de los cuellos de botella legales y políticos. El hecho de comparar los acuerdos entre las infraestructuras de investigación y los prestadores de servicios es una forma de evitar potenciales complicaciones.

GLOBIS-B es un acuerdo entre doce infraestructuras de investigación (Kissling *et al.*, 2015) que se presentan en la siguiente tabla.

Table 1. Project partners and supporting research infrastructures of the GLOBIS-B project. The listed supporting research infrastructures represent those that have agreed to contribute to the GLOBIS-B project.

Acronym	Organisation	Geographic scope	Website
<i>Project partners</i>			
UvA	University of Amsterdam (Institute for Biodiversity and Ecosystem Dynamics)	Netherlands	http://ibed.uva.nl/
CU	Cardiff University (School of Computer Science and Informatics)	UK	http://www.cs.cf.ac.uk/
GNUBILA	gnúbila France	France	https://gnubila.fr/
CNR	Consiglio Nazionale delle Ricerche (Institute of Biomembranes and Bioenergetics)	Italy	http://www.cnr.it/sitocnr/home.html
FI-UAH	Universidad de Alcalá (Instituto Benjamin Franklin)	Spain	http://www.institutofranklin.net/
MLU	Martin-Luther-Universität Halle-Wittenberg (German Centre for Integrative Biodiversity Research i-Div)	Germany	http://www.idiv-biodiversity.de/idiv/research/geo-bon/
<i>Supporting research infrastructures</i>			
Atlas	Atlas of Living Australia	Australia	http://www.ala.org.au/
BC-CAS	Biodiversity Committee of the Chinese Academy of Sciences	China	http://www.kepingma.com/index.html
CRIA	Brazilian Reference Centre on Environmental Information	Brazil	http://www.cria.org.br/
DataONE	Data Observation Network for Earth	USA	http://www.dataone.org/
ELIXIR	European infrastructure for biological information	Europe	http://www.elixir-europe.org/
GBIF	Global Biodiversity Information Facility	Global	http://www.gbif.org/
GEO	Group on Earth Observations Biodiversity Observation Network	Global	http://www.geobon.org
BON			
GBoWS	Germplasm Bank of Wild Species at Kunming Institute of Botany	China	http://english.kib.cas.cn/
LifeWatch	European Infrastructure for Biodiversity and Ecosystem Research	Europe	http://lifewatch.eu/
NEON	National Ecological Observatory Network	USA	http://www.neoninc.org/
SANBI	South African National Biodiversity Institute	South Africa	http://www.sanbi.org/
WDCM	World Data Centre of Microorganisms at WFCC-MIRCEN	Global	http://www.wdcm.org/

Figura 108. E-infraestructuras de investigación especializadas en datos de biodiversidad. Fuente: Alonso, 2016a.

Su objetivo principal se encuentra en la movilización de datos y la producción de conocimiento. GLOBIS-B persigue facilitar una cooperación multilateral entre estas infraestructuras de investigación con la finalidad de avanzar en la implementación de las EBV.

Para ello, deberán: (a) especificar los requerimientos de los usuarios para extraer, gestionar, visualizar y analizar los datos de biodiversidad exigidos, (b) identificar de qué manera las infraestructuras de investigación pueden potenciar las capacidades existentes o desarrollar otras nuevas, (c) alcanzar acuerdos en soluciones realistas para dar soporte a los requerimientos de los usuarios de manera que las infraestructuras de investigación puedan ofrecer servicios específicos para calcular

las EBV seleccionadas, (d) diseñar las mejores prácticas para el soporte de las infraestructuras y (e) gestionar las implicaciones legales en lo que respecta a las licencias, derechos de propiedad intelectual o *copyright* y a la compartición de recursos.

El progreso en la definición de las EBV se ha visto obstaculizado por la falta de claridad existente en la definición de qué son y, como señalan Schmeller *et al.* (2017), es urgente aclararlo. Lo que estos autores defienden es que una EBV es una variable de estado biológica que es posible medir en determinados puntos temporales y espaciales con la finalidad de documentar un cambio biológico. La EBV observa desde un punto de vista multidimensional y consiste en una serie de mediciones o estimaciones de aspectos esenciales de biodiversidad.

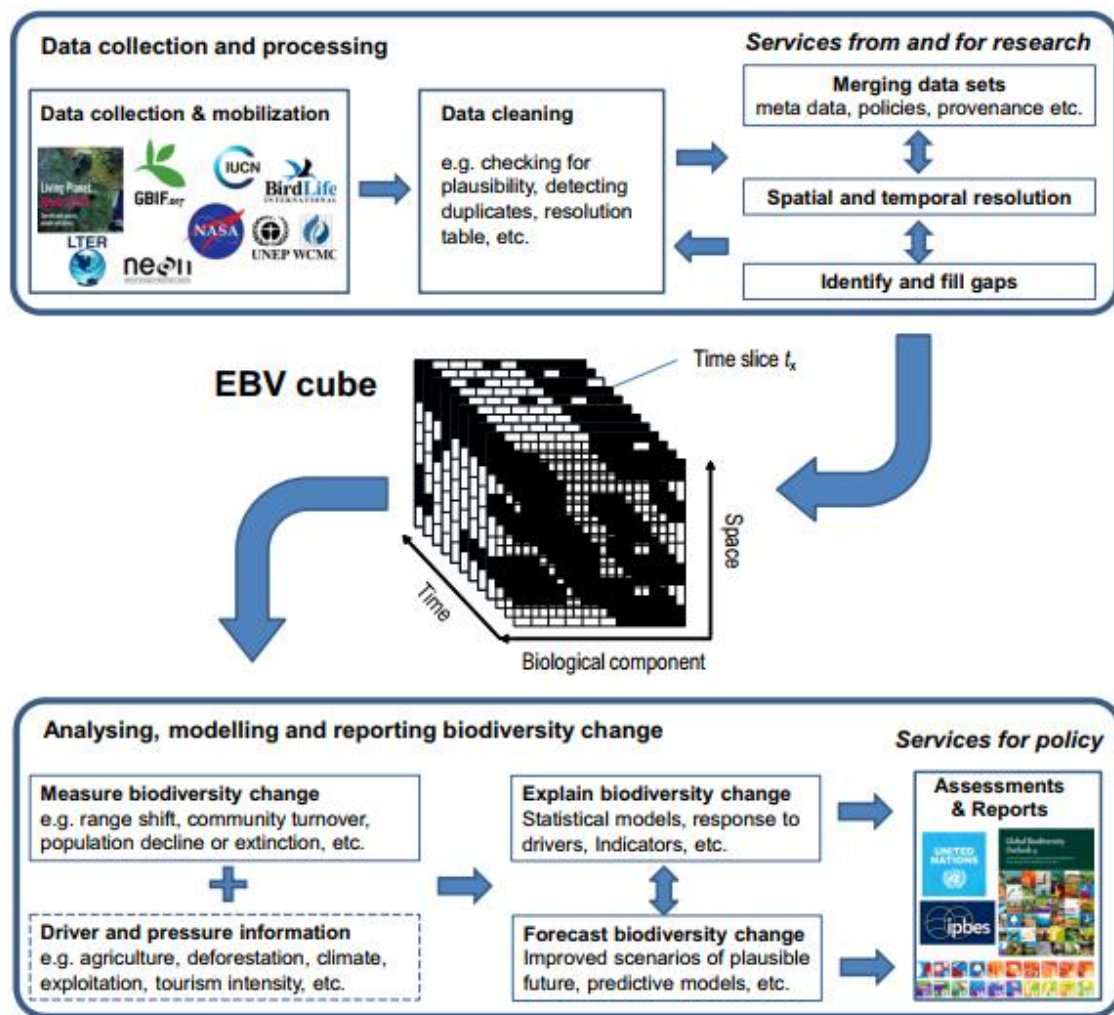


Figura 109. Proceso para que las variables esenciales de biodiversidad sean operativas. Fuente: Schmeller *et al.*, 2017.

Cumplimentar el cubo de EBV requiere recabar datos, movilizarlos y procesarlos como un servicio desde y para la investigación. Una vez que el cubo de EBV tiene

los datos, puede facilitar la medición y la modelización de los cambios de biodiversidad y, por ende, dar servicio tanto a la ciencia como a la política. La totalidad del proceso que arranca de los datos en bruto hasta llegar al cubo de EBV presta también un importante servicio a la comunidad científica, lleva a cabo más análisis eficientes de recursos en lo que afecta al cambio de la biodiversidad en amplios espectros espaciales. Ello incluye cuantificar los *drivers* que subyacen y las presiones en lo que respecta a los cambios de la biodiversidad en los servicios a los ecosistemas (Oliver *et al.*, 2015). Dichos análisis, con validaciones periódicas, servirán, por lo tanto, para alimentar los procesos de creación de políticas regionales o globales que expliquen los cambios observados de biodiversidad, para mejorar las previsiones de cambios de biodiversidad y para escribir informes globales de asesoramiento (Schmeller *et al.*, 2017).

Además, las EBV son clave en la medida en que se eligen para permitir capturar los cambios en la biodiversidad en un rango concreto de tiempo y de espacio. Por eso, es esencial empezar estandarizando, armonizando e integrando datos brutos de investigaciones.

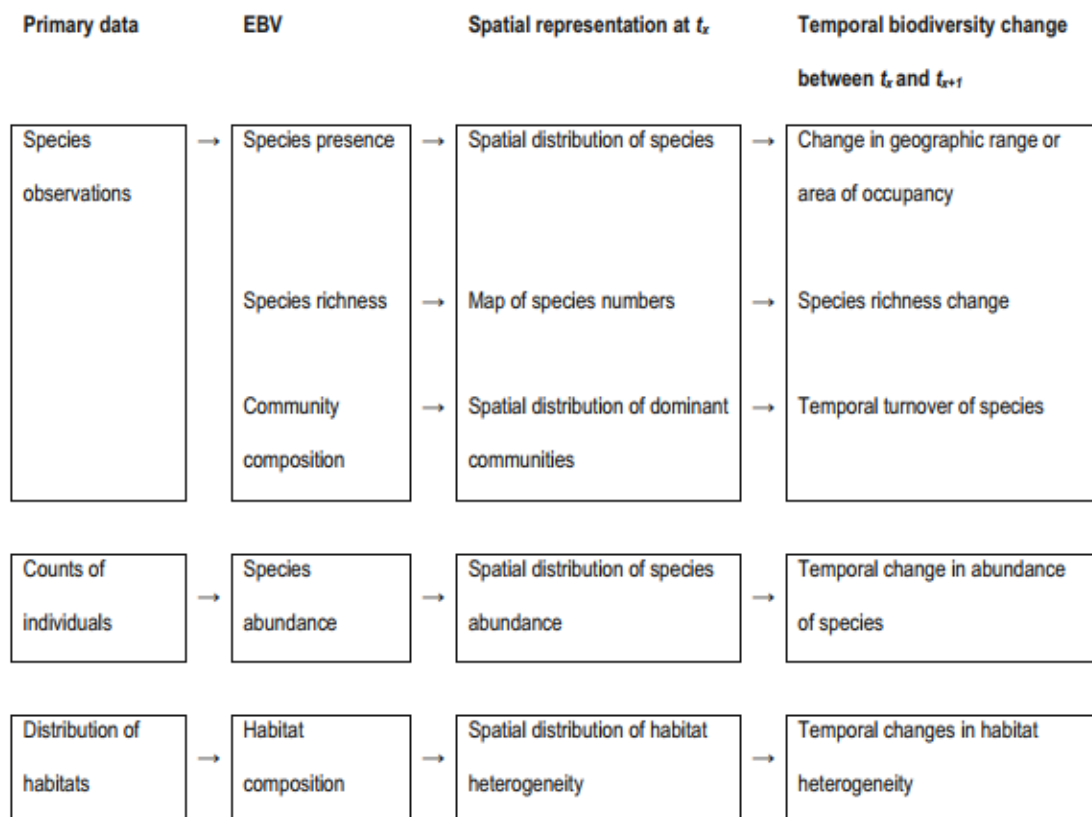


Figura 110. Ejemplos de EBV potenciales, representación en el espacio durante un tiempo determinado y forma en la que pueden usarse para documentar cambios temporales de biodiversidad a la hora de comparar distintas representaciones espaciales en representaciones temporales. Fuente: Schmeller *et al.*, 2017.

Por ejemplo, las cuentas de estos rangos temporales en un momento dado es un rango de la abundancia de especies EBV. Una vez que varios de estos rangos temporales están disponibles, el cambio temporal en la abundancia de especies en el espacio puede documentarse. Los análisis que combinan los datos disponibles en *drivers* y presiones y el cambio observado pueden, en tal caso, servir para informar de la esfera política de las causas que subyacen en este cambio (Schmeller *et al.*, 2017). Por eso, refinar el concepto de EBV es esencial en la medida en que representa un paso más en la operacionalización de las EBV.

De hecho, las EBV son parte de la información de la cadena. Están estratégicamente situadas entre los datos en brotu y los indicadores (Kissling *et al.* 2017).

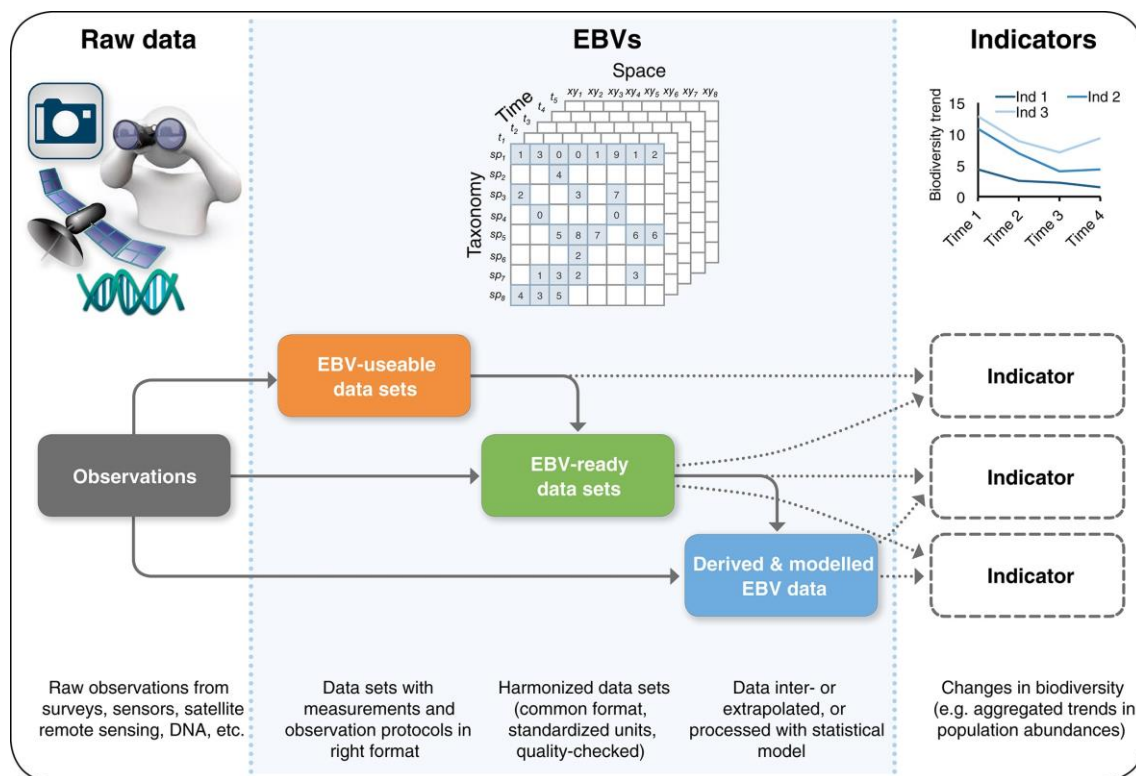


Figura 111. Las EBV forman parte de la cadena de información. Están conceptualmente colocadas entre los datos en bruto y los indicadores. Fuente: Kissling *et al.* 2017.

2. LAS CUESTIONES PRÁCTICAS QUE PLANTEA LA INTEROPERABILIDAD LEGAL EN EL ÁMBITO DE LOS DATOS DE BIODIVERSIDAD

2.1. La interoperabilidad legal en el ámbito del CREaTIVE-B

La última interoperabilidad analizada en el marco de CREaTIVE-B es la interoperabilidad legal. Sobre este punto, se hizo un análisis previo sobre las condiciones de constitución de las diferentes estructuras para identificar las posibles barreras legales que pudieran existir para poder construir una red entre todas. El intercambio de datos y de herramientas con distintas procedencias de autoría y de propiedad exige cuidadosos y eficientes acuerdos en los supuestos en las que las e-infraestructuras involucradas quieran que sus usuarios puedan beneficiarse de los recursos de todas las demás. Esto deviene todavía más importante si se tiene en cuenta la automatización del procesamiento de datos y las interacciones de autorizaciones «máquina-máquina».

Aunque, en un principio, podía plantear muchas dudas, resultó que no había muchas limitaciones desde el punto de vista jurídico para facilitar la interacción de todas las e-infraestructuras. No existían restricciones a la hora de adoptar estándares y protocolos. Por lo tanto, las e-infraestructuras tenían la libertad de modificarlos para adoptar estándares y protocolos comunes.⁸⁵

Por consiguiente, sigue el *roadmap*, las e-infraestructuras llegaron a acuerdos para trabajar y compartir los estándares y protocolos de modo que pudiera analizarse su uso, así como las cláusulas contractuales que usan con sus usuarios. En lo que respecta al *software* en el que están desarrolladas, las e-infraestructuras trabajan fundamentalmente en entornos de código abierto.

Todas las e-infraestructuras trabajan sin restricciones en lo que respecta a la reutilización de datos.⁸⁶ Aunque lo hacen basándose en el acceso abierto de datos, todas las e-infraestructuras tienen una política pragmática en lo que respecta a la reclamación sobre propiedad de datos, separándose del síndrome *it's my data*. Ciertamente, el *roadmap* reconoce que pueden existir restricciones mayores de la interacción con proveedores externos de datos. Otra de las cuestiones relevantes es el posible impacto que pueda tener en términos de acceso la nueva directiva europea de minería de datos, que podría exigir el uso de licencias para acceder a las búsquedas de datos. Si bien es cierto que, como ya se ha puesto de manifiesto,

⁸⁵ Con una salvedad, la exigencia de la Directiva 2007/2/EC, INSPIRE en lo que respecta a la exigencia de usar el sistema GIS.

⁸⁶ Con las salvedades de lo previsto en el artículo 15 del Convenio de Diversidad Biológica en particular, en lo que respecta a Brasil y Sudáfrica.

parece que van a incluirse exenciones en relación con los datos científicos o de investigación.

En este caso concreto, se observa que, como consecuencia de los derechos de propiedad intelectual o *copyright*, pueden producirse restricciones en lo que respecta a la interoperabilidad (legal) automática o semiautomática. Estas restricciones pueden tener, también, su origen en desarrollos de *software* o tecnológicos de cualquier otro tipo, así como en las cuestiones de propiedad de los datos. Las distintas interpretaciones de las normas de derechos de propiedad intelectual o *copyright* que existen en el mundo globalizado actualmente en construcción pueden generar una multiplicidad de problemas. Sobre esto se abundará después, al tratar el tema de la interoperabilidad.

En lo que se refiere a las cuestiones jurídicas que se plantean, el *roadmap* enuncia una serie de acciones como pueden ser:

- Abstenerse de prácticas y políticas que dificulten la reutilización de los datos desde el momento en que se requiera que concurran determinadas exigencias para los datos de carácter agregado.
- Promover técnicas de gestión de la calidad de los datos y compartir buenas soluciones.
- Tener en cuenta las nuevas licencias de Creative Commons, CC4 y CC0, cuya finalidad es aportar alternativas a la distribución de datos y facilitar la divulgación de datos científicos o de investigación (Appleyard, 2016).
- Estudiar las implicaciones de los Earth Science Information Partners (ESIP), el protocolo de citación COOPEUS y los estándares de citación de datos de GEOSS.
- Desarrollar soluciones inteligentes para aplicar exenciones de cualquier derecho sobre los datos de cada una de las e-infraestructuras de manera que soporte el procesamiento automático de datos.

2.2. La interoperabilidad legal en el ámbito del GLOBIS-B

Las cuestiones jurídicas de interoperabilidad se analizan en el Deliverable 4 de GLOBIS-B (Alonso García, 2016). El proyecto de GLOBIS-B define su objetivo como facilitar la cooperación multilateral entre infraestructuras de investigación con el objeto de avanzar en la implantación de las EBV.

Los datos científicos son un conjunto particular de datos para toda una serie de cuestiones. Tradicionalmente, estas cuestiones han sido abordadas de diferentes maneras por las comunidades científicas. Algunos de los conjuntos de datos tienen

marcos de acceso y uso muy abierto, incluida la sociedad en general; otros, en cambio, solo están abiertos a miembros registrados de la comunidad y, los últimos se utilizan dentro de sistemas normativos en los que se considera que los datos son propiedad del titular (bien de aquel que los descubre, bien del primero en colocarlos en algún soporte). Aunque la investigación financiada públicamente se mueve de forma progresiva hacia el acceso abierto, todavía queda un largo camino por recorrer (Alonso García, 2016a). En el *Deliverable* se listan las cuestiones fundamentales que se van a abordar en lo que respecta a la interoperabilidad legal en el marco de GLOBIS-B.

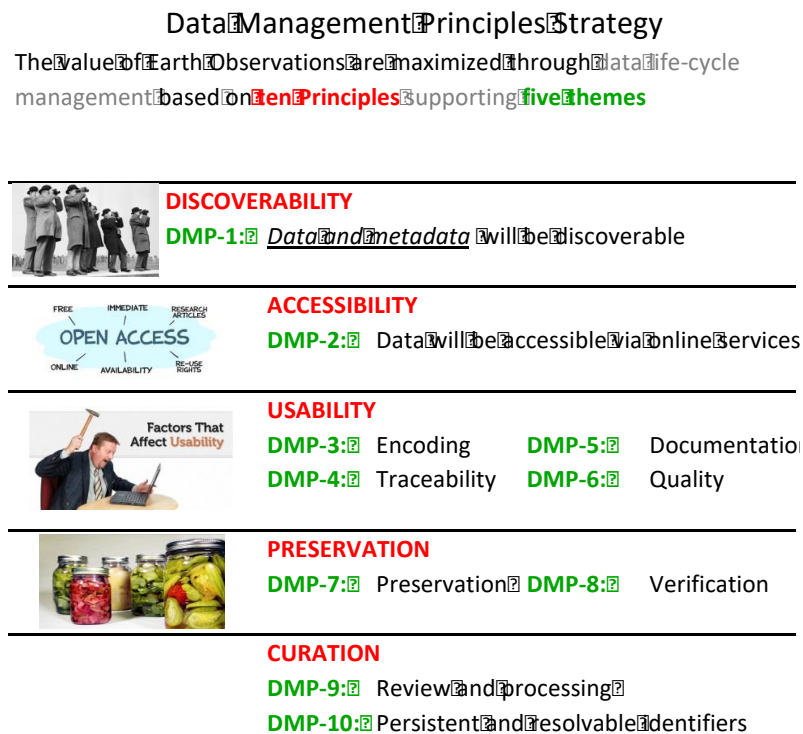
1. ¿Existen enfoques de arriba abajo o de abajo hacia arriba para políticas armonizadas (o mandatos) sobre el acceso abierto y el intercambio de datos de todos los tipos de datos necesarios para medir las EBV (incluyendo cambios temporales y espaciales) y, en particular, en lo que afecta a la abundancia de especies? También es relevante el rol que desempeñan el GBIF y otros conjuntos de datos internacionales, así como los inventarios nacionales de diversidad biológica como fuentes armonizadas adecuadas de datos.
2. Normas y protocolos comunes requeridos por instrumentos jurídicos; bien establecidos en normas legales (nacionales, supranacionales o tratados internacionales), bien a través de las condiciones exigidas en la investigación financiada con fondos públicos o mediante acuerdos contractuales de las diferentes comunidades científicas y de los servicios web de IT que vinculan a proveedores y usuarios de datos. Los efectos de los estándares recomendados v. los totalmente abiertos (por ejemplo, el Marco Común para Datos de Observación de la Tierra de USGEO sobre Observaciones Terrestres (USGEO) contiene una lista de estándares recomendados).
3. Nociones comunes de lo que implica la dedicación de los datos de biodiversidad al dominio público. Evaluación de cuántos conjuntos de datos y de cuáles de ellos están en el dominio público en todo el mundo.
4. Limitaciones legítimas a la hora de imponer restricciones en el acceso abierto de los datos de biodiversidad y del intercambio de datos en general (por ejemplo: en lo que respecta a los recursos genéticos, la localización de especies en peligro de extinción, etc.).
5. En particular, hay que tener en cuenta los límites basados en los distintos ámbitos de la biodiversidad: en primer lugar, la confidencialidad de los datos y los secretos comerciales, así como los periodos de embargo; en segundo lugar, el síndrome *it's my data* y las carreras profesionales basadas en el secreto de los datos. Existen diferentes filosofías de las distintas

comunidades científicas con relevancia para la suministrar los datos necesarios en la medición del cambio de biodiversidad, en particular de las EBV y, más concretamente de la abundancia de especies.

6. El uso de licencias de Creative Commons por parte de los investigadores de biodiversidad y sus límites (en particular, en lo que respecta a las licencias CC0 en relación con los datos).
7. La gestión de las expresiones de los derechos de propiedad intelectual o *copyright* de los datos y de los propios datos en las bases de datos que existen y en los servicios web.
8. Gestión de la atribución de datos relacionados con la biodiversidad y gestión de la agregación de atribuciones en datos digitalizados de biodiversidad.
9. Metadatos que abordan cuestiones en materia de derechos. ¿Resulta útil este enfoque?
10. Límites de las licencias de *software* a la hora de obtener datos o de hacer uso de ellos, en particular, en relación con el *software* GIS y los entornos de *software* actuales.
11. Los datos de biodiversidad tienen un origen científico múltiple, ya que tanto los datos bióticos como los abióticos son esenciales para entender la biodiversidad a todos los niveles (desde el nivel genético hasta el paisaje). La biodiversidad, en esencia, es multidisciplinar. ¿Deberían utilizarse normas (en lugar de reglas) para facilitar el intercambio de datos entre las diferentes comunidades diferenciadas? ¿Deben los servicios web y las infraestructuras electrónicas de la biodiversidad funcionar sobre una base normativa (y no legal)? ¿Es realista crear marcos operativos «alegales» en los que los derechos de propiedad intelectual o *copyright* y otros límites al acceso y uso de datos se hagan a un lado simplemente apartados mediante esquemas normativos basados en una «cultura científica» común?
12. En la era de la economía del conocimiento basada en tecnologías digitalizadas, los datos se han convertido en un valioso activo comercial que puede ser explotado mediante licencias o incluso a través de la venta a terceros. La protección de datos debida al aumento del valor de los datos como un objeto que puede ser el centro de las transacciones *per se* o como una consideración auxiliar de acuerdos de servicios tecnológicos no solo ha estimulado un movimiento legal para aumentar la protección de datos, sino toda una cultura legal de datos, es decir, como un servicio de negocios sujeto a sus propias reglas particulares. El hecho de que el acceso abierto se

imponga como el principio jurídico contrario, debido a la necesidad de avanzar en la ciencia y la política pública (cuando los datos se generan en el contexto de la financiación pública), no puede ocultar que esta cultura jurídica sobre cómo gestionar los datos puede ser aplicable con independencia de que se gestione como un servicio público o como una empresa privada (o sin ánimo de lucro). Los «acuerdos de licencia de datos» representan un subsector de la ley digital de IT que opera bajo un paraguas de ley, costumbre o jurisprudencia que rige diferentes aspectos del proceso. Por lo tanto, no solo existen nuevas normas que regulan la «propiedad de los datos» o «la concesión de licencias de datos». Entre estas, la principal diferencia es que el primer conjunto de normas es ejecutable contra terceros, mientras que las últimas solo son exigibles frente a las otras partes contratantes. La mera interacción con un cliente web puede crear miles de millones de contratos al año, ya sea por meros clics o incluso accediendo a una página web a través de Internet, como es el caso de los llamados acuerdos de *clickwrap* o *browsewrap*. ¿Están los servicios web relacionados con la biodiversidad o las e-infraestructuras condicionadas por tales reglas?

Todas estas cuestiones se han ido tratando con carácter general, a lo largo de este trabajo. Se han abordado en el GEO, como ya se ha visto con anterioridad, dando lugar a dos conjuntos de principios los *Data Management Principles* y los *Data Sharing Principles* (Alonso García, 2016).



DMPs Implementation Guidelines

Description/Explanation of each principle and implementation

Guidance on implementation with examples

Resources for implementing each principle

Metrics measuring adherence to implement each principle

Figura 112. Estrategia de *Data Management Principles*. Fuente: Alonso García, 2016a.

En la RDA se da un paso más con los principios y las pautas de implementación. GLOBIS-B las aborda en el marco definido del proyecto y enfocándose en los datos de biodiversidad. De hecho, la tarea fundamental de los juristas sería señalar en qué puntos las soluciones que dan los científicos y técnicos para medir las EBV se pueden ver debilitadas o pueden encontrarse con limitaciones teniendo en cuenta este tipo de entornos legales o en qué entornos legales equivalentes sería necesario gestionar los datos para que eso funcione.

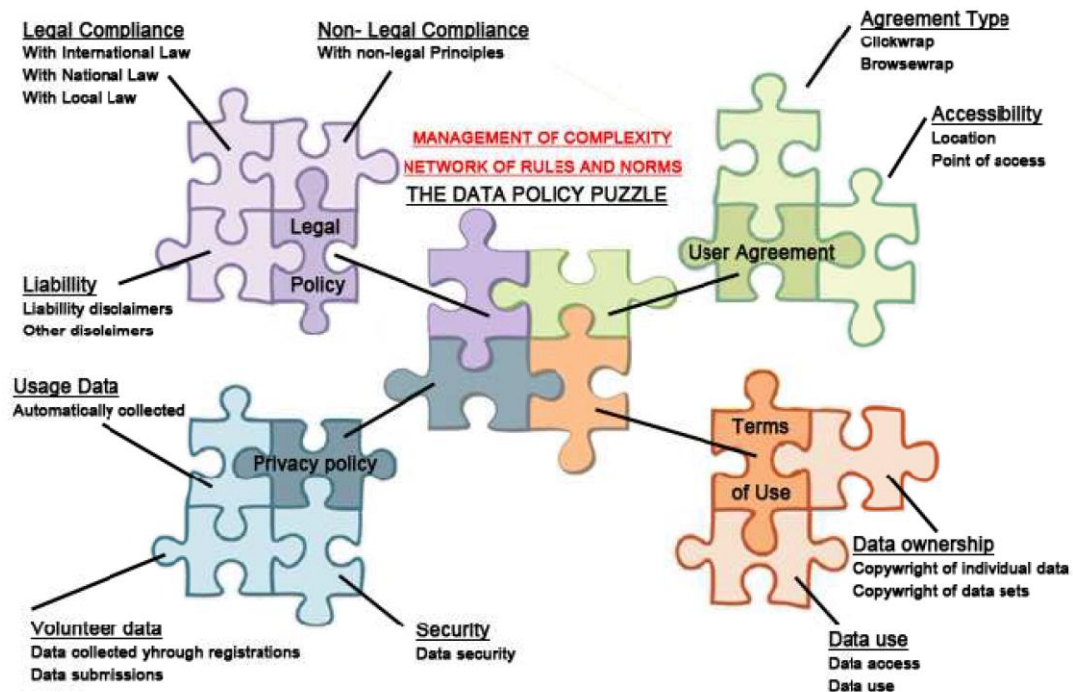


Figura 113. Políticas de datos para la participación pública en la investigación científica elaborado por Anne Bowser. Fuente: Alonso García, 2016a.

La interoperabilidad legal no es un tema resuelto aún. Queda mucho camino por hacer, aunque parece que los pilares se están estableciendo ya.

Un paso más ha dado GLOBIS-B al analizar la interoperabilidad legal en los flujos de datos de las EBV. Además del análisis, establecen una serie de requerimientos:

- 1) Que las condiciones legales de los conjuntos de datos estén claramente definidas y sean claramente identificables para cada uno de ellos,
- 2) Que las condiciones legales de cada conjunto de datos permitan, en cualquier caso, tanto la creación como el uso y combinación de productos derivados,
- 3) Que los usuarios puedan acceder legalmente y hacer uso de cada uno de los conjuntos de datos, sin necesidad de solicitar autorización caso por caso para hacerlo.

En efecto, la interoperabilidad legal es clave para automatizar los flujos de datos y para que pueden desarrollarse con éxito datos y productos relacionados con las EBVs (Kisling *et al.* 2017). De alguna manera, esto está muy relacionado con los ya analizados Fair Principles y la capacidad de hacer estos datos: *findable*, *accesible*, *interoperand* and *reusable*. Es evidente que cualquier limitación en materia de

interoperabilidad legal genera restricciones en el uso, modificación e intercambio de los datos. Estas restricciones son, sin duda, un problema para construir las EBV por eso se ha ido promoviendo, en paralelo, la creación de mecanismos que permitan combinar datos de distintas fuentes con las mínimas restricciones posibles (RDA-CODATA Legal Interoperability Interest Group 2016).

De igual manera, la creación de las EBV requiere la creación de estándares que faciliten la interoperabilidad (Kisling *et al.* 2017).

Conclusiones

Una vez realizado el estudio, se va a proceder a dar respuesta a las cuestiones que se plantearon en la introducción, siguiendo el orden en el que se formularon.

DOS MARCOS DE CONTEXTUALIZACIÓN

A. Primer marco contextual: la revolución digital

A.1. La aparición de las nuevas tecnologías y la creación de un nuevo contexto: **Internet**

Sus antecedentes pueden encontrarse en la Guerra Fría. Estados Unidos, preocupado por la vulnerabilidad del sistema de comunicaciones, promovió el desarrollo de un proyecto de investigación en redes de conmutación de paquetes. Así nació el denominado proyecto ARPA, como un proyecto militar con el objetivo de garantizar las comunicaciones incluso en caso de destrucción parcial de la red; era un sistema muy fiable, con fácil distribución. Permitía utilizar técnicas de compresión con mayor capacidad de transmisión y de encriptado de los datos, de manera que mejoraba su confidencialidad.

El sistema de comunicación fue desarrollándose y empezó a aplicarse con fines académicos y de investigación. Apareció así MILNET. A mediados de los ochenta la National Science Foundation creó su propia red informática y absorbió ARPANET. Con este proyecto, se llevó a cabo un gran desarrollo de las redes. Más adelante, en los años noventa, en el Centro Europeo de Investigaciones Nucleares (CERN), Tim Berners-Lee desarrolló un proyecto de almacenamiento y recuperación de datos, basado en una idea de Ted Nelson en relación con los hipervínculos.

Nació la *World Wide Web* o telaraña mundial, es decir, la red de redes, Internet tal como se conoce hoy. Se creó, con ello, un único espacio, accesible desde cualquier lugar del mundo en el que almacenar, consultar, actualizar y compartir esa información de forma masiva. Este nuevo sistema vincula la información de forma lógica usando las redes.

Internet no es un concepto unívoco. Hace referencia tanto al contenido como al continente. Su definición ha ido evolucionando con los años y su aparición ha generado un cambio de tal entidad que ya se habla de un cambio de paradigma. Con Internet ha llegado la revolución digital y, con ella, la llamada era del conocimiento, sociedad del conocimiento, economía del conocimiento, etc. De la era del *Homo sapiens* se ha pasado a la era del *Homo digitalis*. Este nuevo paradigma ha permeado a todos los niveles: economía, relaciones sociales, política, etc.

A pesar de los matices que las distintas regulaciones han ido introduciendo, es indiscutible que Internet nació con una vocación clara de compartir información, conocimiento, datos, etc. Y su existencia ha multiplicado exponencialmente la capacidad de crear y gestionar conocimiento. Poner límites en Internet es «como querer poner puertas al campo»; no obstante, el hecho de que la red facilite que los datos «estén», no es suficiente. Es, además, necesario que la información esté ordenada, sea accesible, fiable, etc., es decir, que pueda encontrarse. En efecto, con la proliferación de información nace el llamado *big data*.

A.2. La gestión del conocimiento

Así, con la llegada de la tecnología se ha producido una revolución en la forma en que se desarrolla y divulga el conocimiento, y la sociedad está en pleno proceso de transformación. Con esta revolución digital los datos se multiplican de forma exponencial. De hecho, en los últimos cien años el conocimiento ha avanzado mucho más rápido que en los mil novecientos años anteriores.

Hoy el conocimiento está al alcance de todos. El mero acceso a Internet permite el llegar a contenidos que antes estaba disponibles solo de forma limitada, en bibliotecas o en centros de información. Si antes el conocimiento era individual, hoy se está construyendo el concepto de conocimiento colectivo, colaborativo, cuya puesta en común acelera más si cabe la generación de más conocimiento. Aunque recientemente, la desaparición de la neutralidad de la red podrá distorsionar el desarrollo del conocimiento colaborativo tal como lo conocemos hasta hoy. Este concepto se analizará más adelante, cuando se vean las modificaciones habidas en la era Trump.

Este cambio de paradigma está teniendo un impacto brutal en la sociedad. Se dice que más del 40% de las profesiones que existen hoy desaparecerán en los próximos años y que aparecerán otras nuevas —fundamentalmente ligadas a la automatización— que todavía ni existen en nuestra cabeza. La Universidad de Oxford habla incluso de la desaparición de un 47% de las profesiones actuales.

Esto no es un futurible, pues ya hoy en el mercado actual de trabajo existen posiciones disponibles que no llegan a ser cubiertas por falta de profesionales especializados, a pesar de la crisis en la que hemos estado inmersos y los niveles de

paro en los que nos hemos encontrado. Las personas que se encuentran en el mercado de trabajo no están preparadas para asumir las posiciones hoy vacantes. La expectativa es que este problema se acentúe en los próximos años.

De los primeros pasos, a mediados del siglo XX, en la gestión de contenidos mediante el uso de tecnologías entonces «avanzadas» como el *data warehouse* a la gestión de datos dinámicos como se conoce hoy se ha producido un gran salto. El *data warehouse*, como su propio nombre indica, era un almacén de datos y, por lo tanto, tenía una connotación de información estática. En su momento, supuso también un gran cambio favoreciendo la creación de herramientas de CRM (*customer relationship management*). Fueron las primeras herramientas de análisis de datos, que permitían elaborar perfiles de consumidores en función de sus preferencias de consumo. Apareció, entonces, la minería de datos como una forma de explotar esos datos.

Tal vez pueda considerarse precursor de las actuales infraestructuras de gestión del conocimiento. La velocidad y volumen de generación de conocimiento o datos, su crecimiento exponencial, ha hecho que de las primeras compilaciones novedosas, aunque estáticas, de *data warehouse*, se haya pasado a lo que hoy se conoce como *big data*. Un contexto en el que los datos se extraen de una multiplicidad de fuentes cambiantes y dinámicas, que permiten realizar análisis más completos y con mejor fundamento.

Sin embargo, el salto del *big data* no tiene parangón. Afecta a todos los ámbitos. La gestión de los datos deviene clave para la toma de decisiones informadas, casi en tiempo real. Entre 2005 y 2020, según un estudio de IDC, el factor digital se multiplicará por trescientos, de ciento treinta exabits se pasará a cuarenta mil o cuarenta millones de gigabits. A partir de 2020, la expectativa es que se multiplique por dos cada dos años. Se espera, como consecuencia de ello, la aparición de supercomputadores mucho más potentes para trabajar estos datos tanto en volumen como en velocidad.

Hoy, solo un 25% de los datos obtenidos se usan con fines analíticos, mientras que la previsión es que para 2020 este porcentaje sea del 33 % y después de esa fecha, del 40 %. En este contexto, la interconexión es crítica. Se espera que entre 2025 y 2040, llegue la era de la singularidad en la que los seres humanos y las máquinas convivirán y el nivel de automatización será muy alto.

A.3. Regulación de las infraestructuras del conocimiento: diferencias entre la Unión Europea y Estados Unidos

Por eso, es clave hoy clasificar y ordenar el conocimiento para que no se pierda, garantizando su sostenibilidad. Sin duda, es un cambio que exige reformular las herramientas existentes hasta este momento. Así, una de las primeras cuestiones que

deben abordarse es definir de qué manera se gestiona ese conocimiento. De pronto, las formas de gestión del conocimiento parecen haber quedado obsoletas para los volúmenes gestionados hoy y para las formas de darles acceso. Por consiguiente, se ha invertido el problema: el exceso de contenidos y de información puede hacer que el conocimiento relevante quede «hundido» y pase desapercibido. Se ha acuñado un nuevo término: la «infoxicación», un nuevo concepto nacido de la intoxicación que produce el exceso de información no sistematizada. Consecuencia de todo esto es que uno de los problemas fundamentales sea ordenar, sistematizar, etc.

A.3.1. Las infraestructuras del conocimiento

Como respuesta a esta cuestión, aparecen las llamadas infraestructuras del conocimiento, ciberinfraestructuras, e-infraestructuras, etc. Aunque existe una multiplicidad de denominaciones, el objetivo es único: conseguir, almacenar, clasificar y gestionar conocimiento en Internet, tal vez, primero, reproduciendo lo que hacían las bibliotecas antes, es decir mediante la creación de repositorios de acceso electrónico

La creación de infraestructuras del conocimiento, repositorios estáticos al principio, es un primer paso, aunque no es suficiente. Debe promoverse, facilitarse y, además, que dichas infraestructuras puedan interactuar, conectarse entre sí y convertirse en fuentes de conocimiento mucho más completas. Por eso, es clave la interdependencia entre ellas, cada vez más frecuente e importante.

No solo sirven para almacenar, sistematizar y estructurar el conocimiento, sino que es también fundamental que lo hagan accesible y, en última instancia, reutilizable. Aunque ya antes los usuarios eran, simultáneamente, consumidores y productores de conocimiento, la revolución digital ha multiplicado la incidencia y la velocidad a la que esto se produce.

La accesibilidad al conocimiento tiene, por consiguiente, un efecto multiplicador. Así, el reto principal de las e-infraestructuras no es solo la gestión de los datos y de su volumen, sino la capacidad de generar valor a través de ellos o su análisis para poder tomar decisiones..., casi en tiempo real. Para ello, hay que estar seguro de que el conocimiento, es decir los datos, son confiables. De alguna manera, poder comprobar que es consistente y que los datos obtenidos pueden reproducirse en el tiempo.

También es necesario verificar la integridad del conocimiento. Es decir, identificar las herramientas que se usan para la elaboración de infraestructuras e ir trasladando el conocimiento, según avance la tecnología. Si la infraestructura quedase obsoleta, podría perderse el contenido. Por último, es necesario que el sistema sea sostenible a

medio y largo plazo. Para ello, se hace imprescindible buscar métodos sostenibles que no limiten el acceso a los datos y al conocimiento.

A.3.2. La aparición de la e-ciencia

De la aplicación de los nuevos paradigmas del *big data* a la ciencia tradicional aparece la e-ciencia, lo que muestra que también la ciencia está en proceso de evolución. En un primer momento, la ciencia se limitaba a la observación empírica de los fenómenos naturales; después, dio un paso, empezando a teorizar sobre la observación empírica. Construyó, así, los primeros modelos y generalizaciones. En las últimas décadas, la ciencia ha añadido un tercer paradigma: la simulación informática. Hoy se ha dado un cuarto paso: la gestión de datos que unifica teoría, experimentación y simulación. La ciencia necesita estar preparada para todos los nuevos retos que nos esperan en la era del *big data*.

Procede, a estos efectos, analizar la construcción del concepto de dato en general y del de dato científico o dato de los procesos de investigación, en particular. Más importante que la denominación, en si misma, es la construcción conceptual del mismo. Aunque la construcción conceptual es parecida, existen dos sistemas jurídicos que abordan los conceptos de forma muy distinta. Por un lado, está la construcción europea (donde se usa más el término e-infraestructuras) y, por otro, la estadounidense (que habla, más comúnmente, de ciberinfraestructuras).

A.3.3. Dos enfoques diferentes: e-infraestructuras o ciber-infraestructuras

A) El sistema europeo: la e-infraestructuras

En Europa se dan cuenta de la importancia e impacto de la innovación tecnológica y científica en su economía. De hecho, se habla ya de dependencia. Los problemas globales que es necesario abordar requieren fuertes capacidades de innovación con un acceso de calidad a las infraestructuras de investigación. En el contexto europeo, se utiliza el término e-infraestructura. Estas se convierten en el instrumento clave de encuentro de investigadores, políticos, agencias que subvencionan e industria, con la finalidad de abordar de forma conjunta los distintos problemas actuales.

Como consecuencia de ello, en el año 2002, los Estados miembros y la Comisión Europea decidieron crear el Foro Europeo Estratégico de Infraestructuras de Investigación (ESFRI), con la finalidad de reforzar la integración científica de Europa y su proyección internacional: el entorno necesario para que los Estados pudieran explorar las áreas comunes a fin de desarrollar infraestructuras de investigación con relevancia paneuropea. Para lograrlo, trabajaron conjuntamente en

el desarrollo de un *roadmap*. El primero se publicó en 2006, y se actualizó en 2008 y en 2010.

Este *roadmap* recogió la planificación en relación con todas las disciplinas científicas para el desarrollo de infraestructuras de investigación tanto a escala europea como internacional. Los ESFRI también querían contribuir a reducir la fragmentación promoviendo que los resultados pudieran ponerse a disposición de todos los investigadores europeos. Su misión era servir de base para el desarrollo de un enfoque coherente y estratégico a la realización de políticas, en lo que respecta a las e-infraestructuras europeas y paneuropeas.

De alguna manera, la formalización y publicación de la hoja de ruta de los ESFRI contribuyó a fomentar la colaboración internacional, incluso más allá de la Unión Europea. Como consecuencia de ello, creció la colaboración con otros países: Australia, Estados Unidos, Japón, Sudáfrica, etc. Como se dieron cuenta de que no existían instrumentos legales apropiados para la construcción de dichas infraestructuras de investigación a nivel europeo, en las reuniones de la Comisión Europea empezó a trabajarse en la construcción de un marco legal adecuado.

Uno de los problemas en materia de colaboración internacional era la necesidad de que los acuerdos cristalizaran en tratados internacionales, con la complejidad que eso suponía a la hora de negociar, ratificar, etc. Con la finalidad de simplificar el proceso y dar una respuesta más avanzada, se recurrió a la creación, en la Unión Europea, de una forma jurídica nueva: los denominados «consorcios de infraestructuras de investigación europeas» (ERIC).

Además, la Unión Europea se dio cuenta de que el formato de apoyo vía subvenciones resultaba insuficiente, y buscó implantar un marco jurídico más adecuado para la creación y explotación de infraestructuras de investigación. Su función principal era la creación y explotación de estas sin fines lucrativos. Esta *joint venture* europea, permite, también, la participación de terceros no europeos.

La construcción de estas e-infraestructuras hizo necesario armonizar sistemas jurídicos y culturas diferentes mediante instituciones supranacionales. La «particularidad» de estas infraestructuras, en realidad, era una consecuencia lógica de la propia naturaleza —supranacional— de la Unión Europea y del hecho de que las economías de escala fueran mucho más ventajosas a ese nivel que a nivel nacional, puesto que muchos Estados miembros no podrían permitirse tener estas e-infraestructuras en solitario.

Esta forma jurídica concreta, por consiguiente, simplificaba la creación de infraestructuras de investigación con intereses paneuropeos así como su funcionamiento. Por motivos de transparencia, el ERIC debía crearse mediante la adopción de una decisión, que, a su vez, debía publicarse en el *Diario Oficial de las*

Comunidades Europeas (DOCE). De hecho, los ERIC adquirirían personalidad jurídica (reconocida en todos los países de la UE) desde la fecha en que la Comisión Europea dictase la decisión.

Los ERIC, como estructuras jurídicas, permiten crear nuevas e-infraestructuras, pero también que e-infraestructuras ya existentes decidieran adoptar esta forma jurídica, por considerarla más adecuada a la finalidad buscada. Los miembros de un ERIC adquirirían un compromiso de contribución al mismo y su articulación tuvo su origen en la forma continental de regular en Europa, mediante normas, regulaciones o legislaciones concretas que establecían, de forma cierta, su funcionamiento, las condiciones que tenían que cumplir, etc. En efecto, en Europa hay una cultura predominante de establecer regulaciones de referencia para articular este tipo de organizaciones.

Quedaron, pues, regulados en el Reglamento (CE) n.º 723/2009 del Consejo, de 25 de junio de 2009 (Parlamento Europeo y Consejo de la Unión Europea, 2003) relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas. Existe un número limitado de ERIC en Europa, que no se van a enumerar en las conclusiones, por estarlo ya en el texto principal de la tesis en las páginas 82 y siguientes.

El Reglamento establece que los Estados miembros del Consorcio serán los responsables de los gastos e ingresos. La Comisión no financia los ERIC. Por lo tanto, el primer paso es construir un marco jurídico, una forma de crear instituciones con personalidad jurídica propia, y, por ende, con autonomía para funcionar. Con la finalidad de promover la I+D+i en la Unión Europea, se crean distintos programas marco, programas de inversión en I+D, con medidas incentivadoras de la recepción de los mismos: un proyecto en el que concurren distintos países, con la finalidad de que crear un marco de colaboración entre países, empresas del sector privado, universidades, etc. para que puedan trabajar conjuntamente. Los ERIC tienen la posibilidad de competir, en igualdad de condiciones con otras instituciones nacionales y de cualquier otra índole. La Comisión si que financia las fases preparatorias de los ERIC, para impulsar/facilitar su creación.

La realidad es que los ERIC cuentan con un plus de colaboración entre Estados que les puede servir para facilitarles el acceso a los fondos estructurales, a los FEDER... en todo caso, en régimen de competencia con cualesquiera otras instituciones que quieran acceder a los mismos.

B) El sistema de los Estados Unidos: las ciber-infraestructuras.

En cambio, el sistema de Estados Unidos sigue un modelo completamente distinto del europeo. En este contexto, en Estados Unidos se prefiere el término ciber-infraestructura. No existen marcos de referencia generales ni regulaciones genéricas, sino construcciones y financiaciones *ad hoc*. Existen también una serie de organizaciones que dan subvenciones: agencias federales, como la National Science Foundation, agencias gubernamentales, como el National Institute of Health, y otras organizaciones como universidades o fundaciones sin ánimo de lucro. La concesión de la subvención está condicionada a reunir las condiciones concretas exigidas en cualquier subvención, por parte de cualquier organismo.

La National Science Foundation, una agencia federal independiente, creada en 1950 por el Congreso, tiene como misión contribuir al progreso de la ciencia, promover el bienestar, la prosperidad y la salud nacional y garantizar la defensa nacional. Es una agencia fundamental porque subvenciona la investigación básica. Considera esta subvención clave, en la medida en que es uno de los principales motores de la economía estadounidense, y sirve para promover el conocimiento consolidando el liderazgo global.

Esta misma agencia concede subvenciones en todos los ámbitos científicos (geología, zoología, astrología, etc.) y de ingeniería, con la salvedad de la medicina. Opera *bottom-up*, es decir, haciendo un meticuloso seguimiento de la investigación básica dentro de Estados Unidos y poniendo subvenciones a disposición de proyectos concretos que deben aplicar y reunir una serie de requisitos. Subvenciona aproximadamente el 20 % de toda la investigación básica soportada por el Gobierno federal en todos los colegios y universidades estadounidenses.

Otra fuente importante de subvenciones en Estados Unidos es el National Institute of Health (NIH), que forma parte del Departamento Federal de Salud y Servicios Humanos de Estados Unidos. Está integrado por veintisiete centros e institutos con sus propias agendas de investigación. Está centrada en dar subvenciones para investigación médica, tiene una planificación de investigación y subvenciona proyectos ajustados a sus líneas de investigación.

Tanto el tipo de subvención como el órgano del que procede condicionan la posterior utilización y reutilización de los datos. Existe una base de datos general, que permite encontrar todas las subvenciones existentes para la financiación de un proyecto científico o de una investigación. El hecho de que las investigaciones estén subvencionadas por agencias gubernamentales tiene consecuencias prácticas: de toda la investigación financiada, los resultados obtenidos entran en el dominio público, por la aplicación de las llamadas políticas de acceso público, sobre las que se volverá después. Asegurarse de que los datos de investigación (financiada) sean

accesibles con la finalidad de que puedan utilizarse el mayor número de veces posibles es una cuestión de gestión racional de los recursos públicos.

La National Science Foundation exige, además, un plan de gestión de datos de todas las nuevas propuestas con la finalidad de asegurarse de que los investigadores tienen planificado mantener los datos durante la investigación, pero también después, con la finalidad de favorecer el intercambio de datos. El llamado plan de gestión de datos debe describir las fuentes y recursos físicos y virtuales.

Además, dispone de una política de archivo de datos todavía más específica con directrices en función de sus categorías. De igual manera, el National Institute of Health tiene requisitos concretos de estructuras de preservación de los datos de las publicaciones científicas. Todo lo que se genere en este marco de subvenciones deberá publicarse en *PubMed*. Es más, la concesión de subvenciones está condicionada a la presentación de un plan de difusión, con la finalidad de compartir los resultados de los estudios realizados. Toda esta información pasará a formar parte del dominio público, es decir, no tendrá restricciones de uso derivadas de los derechos de propiedad intelectual.

De la necesaria interdependencia de las e-infraestructuras surge toda la problemática de las barreras que dificultan la interacción, lo que se conoce con el nombre de interoperabilidad, tema se abordará enseguida, analizando en particular las cuestiones jurídicas que suscita.

A.4. Las cuestiones jurídicas que plantea la revolución digital: la interoperabilidad

La interoperabilidad es, de partida, un concepto técnico. En efecto, la interoperabilidad es antes que nada una cuestión de índole técnica. Así, la interoperabilidad de datos es la «propiedad de un producto o sistema [...] de trabajar con otros productos o sistemas, actuales o futuros sin ninguna restricción en su acceso o implementación»⁸⁷ (Alonso García, 2016, *b*). Existen, por lo tanto, distintos niveles de interoperabilidad siendo la primera la interoperabilidad técnica. Sin ella, las demás no pueden existir.

⁸⁷ Recuperado de: <http://interoperability-definition.info/en>.

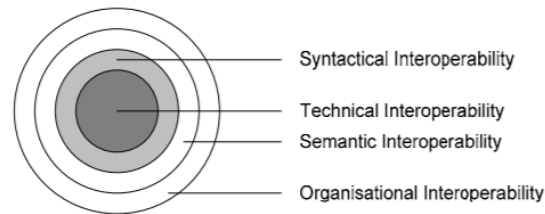


Figure 1: Different levels of interoperability

Figura 114. Distintos niveles de interoperabilidad. Fuente: Van der Veer, 2008.

El siguiente nivel de interoperabilidad es la sintáctica que ocurre cuando dos o más sistemas se pueden comunicar entre sí. Se logra con formatos de datos, protocolos y estándares específicos XML o SQL u otros de menor nivel (Alonso García, 2016*b*). en un sistema de base web, la interoperabilidad sintáctica supone estandarizar la comunicación entre un *software* de cliente y un servidor (Schaeffer *et al.*, 2012).

La Research Data Alliance 2015 define la interoperabilidad semántica como la capacidad de los servicios y de los sistemas de intercambiar datos de una forma útil y con sentido (Sudmanns *et al.*, 2018). Es un concepto mucho más complejo de lo que aparece a primera vista, puesto que requiere que se entienda el sentido de los datos y de la información en el proceso de comunicación (Harvey *et al.* 1999). Con la finalidad de facilitar que exista la interoperabilidad semántica, los *softwares* clientes que solicitan el proceso y los servidores, que llevan a cabo el proceso, acuerdan de forma clara y unívoca el contexto y los conceptos de las entidades y sus relaciones (Sudmanns, 2018). Un ejemplo de problemas con la interoperabilidad semántica y ontológica puede encontrarse en GlobisB donde se analizan lenguajes complejos. La importancia se encuentra en la complejidad que supone el lenguaje y en la falta de armonización de términos y no sólo de que se trate de idiomas distintos. En efecto, un ejemplo podría ser el término *traits* que unas veces se traduce como características y otra como rasgos. La falta de identificación de términos hace que la interoperabilidad semántica se complique.

Sin interoperabilidad, la interacción de los datos es imposible; por lo tanto, el punto de partida de cualquier interoperabilidad es que físicamente exista interoperabilidad técnica, semántica y organizativa. Es más, sin estas interoperabilidades la existencia de los metadatos es imposible. Por eso, son condición previa a la existencia de la interoperabilidad legal.

Acorde con los objetivos de esta tesis, es clave analizar las cuestiones jurídicas que plantea la revolución digital, lo que se conoce como la interoperabilidad legal. Esta arranca del propio concepto y de su construcción. Por eso, interesa aquí, en particular, el concepto de interoperabilidad legal, aunque no pueden dejar de mencionarse los demás tipos, también relevantes.

La interoperabilidad consiste en facilitar la puesta en común de datos con la finalidad de que puedan reutilizarse. Con la aparición de Internet, los límites territoriales se han visto superados, es decir, que los datos circulan a escala mundial, pues la escala nacional se ha quedado pequeña. Sin embargo, las herramientas que proporciona el derecho internacional no son suficientes a día de hoy para abordar cuestiones tan globales. Sobre el concepto de interoperabilidad legal, se volverá un poco más adelante. El eje del concepto de interoperabilidad, en general, se construye sobre el de dato.

El concepto de dato en el ámbito científico plantea muchas cuestiones adicionales. Una de las primeras controversias abiertas entre los científicos, hoy, es precisamente si el término «dato» hace referencia solo a los resultados de la investigación o también a los «datos» (resultantes de la observación) recogidos durante el proceso de elaboración de la observación.

Resulta también controvertido el tema de si, además de los resultados, deben publicarse los datos que han permitido llegar a ellos, de tal manera que pueda profundizarse la investigación o incluso cuestionarse, sin necesidad de volver a partir de cero.

No existe consenso a la hora de definir lo que es un dato, y menos a la hora de definir lo que se considera dato científico o dato en proceso de investigación (*research data*). Subyace el tema de la realidad física pues, a fin de cuentas, el dato científico parte de la observación de la realidad. En lo que respecta a los datos científicos, diferencian distintas categorías: datos de observación, datos de experimentación, datos de simulación y datos catalogados.

A la hora de analizar la interoperabilidad, es clave analizar los obstáculos jurídicos existentes en relación con el dato, dentro de los cuales se plantean dos claves esenciales: las cuestiones relacionadas con su propiedad y las relacionadas con la protección de los mismos.

A.4.1. La apropiabilidad de los datos

En lo que respecta a su propiedad, existe un gran debate doctrinal con dos corrientes diferenciadas. La más clásica se aferra al concepto de propiedad del dato ligado al soporte en el que se encuentra. Sobre este punto existe una cuestión en relación con quién es, en caso de existir dicha propiedad, el propietario real de los datos. Aquí de nuevo los más conservadores consideran que los datos «pertenecen» a los científicos o investigadores, que, interesados en las publicaciones, han tratado de apropiarse de los datos. Sin embargo, parece que va imponiéndose la tesis de que, de existir tal propiedad, esta podría variar en función de quien financie los proyectos (Estados, organizaciones, etc.).

La otra corriente, más moderna, cuestiona el hecho de que los datos puedan pertenecer a alguien: los datos derivados de la investigación subvencionada deben estar en el dominio público. Esta corriente está fundamentada en que la mayoría de las investigaciones se subvencionan con fondos públicos. La problemática de la «apropiabilidad de los datos» afecta tanto a los datos que sirven de base para la investigación (*raw data* o datos en bruto) como a los resultados obtenidos. Desde luego, si existen dudas en lo que respecta a la apropiabilidad de los datos en general, estas se multiplican en el caso de los llamados *raw data*: teniendo en cuenta que son observaciones de la realidad, los científicos se limitan a «descubrir» algo que ya existe. La esencia de la *Research Data Law* es la no susceptibilidad de apropiación ni *ex lege* ni mediante confidencialidad o derecho de propiedad intelectual de base legal o contractual de datos sometidos al estatuto de acceso y reutilización en abierto de todo dato obtenido mediante financiación pública (Alonso García, 2016).

La propiedad de los datos es una forma de obstaculizar la reutilización de los mismos, puesto que permite establecer restricciones para su uso. En efecto, si los datos son susceptibles de apropiación, el hecho de que estén publicados no es suficiente para que puedan reutilizarse en nuevas investigaciones. De hecho, parece que el concepto de propiedad de los datos ha quedado descontextualizado y obsoleto en el mundo actual. Va cobrando más fuerza el de «titular de los datos», entendido como aquel que, en un momento dado, pueda tener algún derecho sobre los mismos.

Con independencia de que los datos sean o no susceptibles de apropiación, se están buscando alternativas que permitan darles acceso y que se reutilicen con toda una serie de herramientas y licencias.

A.4.2. La protección de los datos mediante sistemas de propiedad intelectual

Otra de las cuestiones clave en el análisis de la interoperabilidad es el de si es susceptible o no que existan derechos de propiedad intelectual sobre los datos. A pesar de las discrepancias existentes en relación con esta cuestión, hoy la tendencia preponderante es pensar que los datos de investigación, que proceden de la observación de la realidad, tienen la consideración de hechos (*facts*) y, por ende, no son susceptibles de protegerse con arreglo al derecho de propiedad intelectual o *copyright*.

De unos años a esta parte, se está consolidando una nueva teoría de los derechos de autor cuyo eje central se ha desplazado. Si antes la justificación de la protección de los derechos estaba en la protección de la innovación, parece que ahora se mueve hacia la apertura de los datos, el hecho de compartirlos, con la finalidad de que la

ciencia avance más rápido, lo que hace que entremos en la era colaborativa o de la colaboración (economía colaborativa, etc.).

Por eso, la conclusión es que los datos científicos serán cada vez menos susceptibles de protección, si es que, como apuntan algunos autores, todavía lo son. La limitación de la protección es el futuro, aunque no se va a producir de forma tan inmediata. A pesar de que en algunas disciplinas los científicos están todavía muy enfocados en la protección de los datos, la interacción e interdependencia hacen que ese sistema ya esté obsoleto. Es necesario trabajar en la construcción de un nuevo paradigma que no frene la innovación ni la ciencia.

La protección de las bases de datos varía en función del sistema jurídico desde el que se aborde. Dos sistemas interesan a efectos de este trabajo: el europeo y el estadounidense.

A.4.3. El sistema europeo de protección

Dentro del modelo europeo de protección de los datos existen tres niveles de protección: el primero afecta al contenido de las bases de datos, es decir, a la protección de los datos en sí mismos, o sea, a la protección desde el punto de vista de los derechos de propiedad intelectual, de los datos en sí (por ejemplo: una fotografía de un ejemplar de una especie amenazada).

Aunque los datos pueden ser susceptibles de protección al amparo de las normas de propiedad intelectual, el cuestionamiento, en lo que respecta a los datos de investigación y científicos relacionados con la biodiversidad, es cada vez más fuerte. Al fin y al cabo, una foto que representa un elefante es una representación de la realidad que no debería ser susceptible de apropiación. En la era digital, la excesiva protección de estos datos podría, sin lugar a dudas, bloquear el desarrollo científico.

El segundo nivel de protección afecta al continente, a la estructura de la base de datos que queda protegida por razón de la inversión que supone. De hecho, se discute —con arreglo a los Acuerdos sobre los Aspectos de los Derechos de Propiedad Intelectual (1994) (conocidos por sus siglas ADPIC o, en inglés, TRIPS)— si, por los asuntos de que se trata o por la forma en que están dispuestos, constituye una creación intelectual propia de su autor.

La normativa europea reconoció que las bases de datos eran una creación intelectual del autor. De alguna manera, la directiva recogió el principio de protección conocido en Estados Unidos como *sweat of the brow*, que reconocía la originalidad en la mera inversión (esfuerzo) realizada.

El tercer nivel afecta a la protección articulada a través del llamado derecho *sui generis*, un derecho que no existe en ninguna otra jurisdicción como tal. De acuerdo

con este régimen, el fabricante de una base de datos, persona física o jurídica, puede prohibir la extracción o la reutilización no autorizada del contenido de una base de datos.

Este derecho tiene cuatro características fundamentales: en primer lugar, persigue proteger la inversión sustancial que supone la creación de bases de datos; en segundo lugar, protege contra los actos de extracción y reutilización del contenido; en tercer lugar, excluye de la protección las partes no sustanciales y, por último, prevé excepciones para ciertos usos: uso privado, con fines de enseñanza, científicos, etc.

Con la multiplicación de los datos nacen nuevas formas de acceder a los mismos. Se crea, así, una nueva tecnología que permite hacer búsquedas de datos mediante la utilización de parámetros y explotarlos con distintas finalidades. Esto es lo que se conoce como el *text and data mining* (minería de datos, en español).

Esta herramienta, que ya empezó a usarse con el *data warehouse* (almacenes estáticos de datos), ha cobrado todavía más importancia con el *big data*. Permite realizar análisis de datos y lograr el procesamiento de la información, e incluye la categorización de texto, la extracción de conceptos, el reconocimiento de patrones, etc.

Después de poner de manifiesto el increíble potencial que tiene la minería de texto y de datos y de señalar que es una herramienta científica de investigación muy poderosa para analizar grandes cantidades de texto y de datos —por ejemplo: los supuestos de publicaciones científicas y de conjuntos de datos de investigación—, y amparándose en la falta de seguridad jurídica que, según señala, existe ahora en la reforma de derecho de propiedad intelectual, la Comisión ha propuesto crear un nuevo derecho de propiedad intelectual, una nueva licencia de aplicable al *text and data mining*. Con la finalidad de no bloquear la investigación pretende establecer una serie de exclusiones y excepciones a la aplicación de la licencia en el marco de las organizaciones de investigación. Cabe preguntarse si esto supone un paso atrás. En efecto, la introducción de limitaciones a la hora de hacer búsquedas y minería de datos supondría una vuelta a la edad de piedra. Sería lo mismo que prohibir, en el mundo offline, hacer búsquedas en un libro ya leído. No parece que tenga mucho sentido.

Esto lleva a la reflexión de cómo se van a definir en la directiva estas excepciones. En el mundo de hoy, donde la velocidad es clave, la definición resulta complicada. La aplicación de una directiva de esta índole es más una barrera para el desarrollo del conocimiento que otra cosa. La pregunta pertinente sería si no existen otras formas más adecuadas a la realidad actual que una limitación de esta índole sobre la capacidad de buscar y tratar los datos.

Teniendo en cuenta los problemas y la falta de consenso en la definición de dato en general, y de dato científico o en proceso de investigación en particular, no parece suficiente dicha justificación. Existen formas alternativas para alcanzar la seguridad jurídica que la Comisión persigue sin añadir licencias que dificulten la reutilización de los datos.

Existe también todo un movimiento en contra de la creación de esta licencia. Gravar el procesamiento de los datos, en un entorno en el que se está trabajando para eliminar las barreras a la utilización y reutilización de los datos, no tiene sentido. Aplicar exenciones a las organizaciones de investigación no es suficiente para que no aparezcan barreras nuevas.

La implementación de esta licencia en Internet puede tener un elevado impacto a la hora de utilizar buscadores, etc. La complejidad de discriminar que las búsquedas procedan de organizaciones científicas es muy alta. Además, habrá que regular el tipo de organizaciones científicas. Este sistema puede paralizar o dificultar, *de facto*, la investigación.

Incluso puede ocurrir que, para evitar la aplicación de la norma europea, se busquen otras jurisdicciones que no establezcan restricciones a estos efectos. Los efectos que la aprobación de esta directiva podría tener son imprevisibles: desde la deslocalización fuera de la Unión Europea hasta la creación de una mayor inseguridad jurídica ante la necesidad de analizar las excepciones y exclusiones, entre otros. Por todo lo anterior, el instrumento elegido no parece el adecuado para resolver las cuestiones que preocupan a la Comisión. Muy relacionado con esto está también la ya mencionada desaparición de la neutralidad de la red, recientemente aprobada, y cuyas consecuencias, hoy, son imprevisibles. Si bien existe una enorme preocupación de como va a impactar a la circulación y creación de contenidos.

A.4.4. El sistema de protección de Estados Unidos

A diferencia de lo que ocurre en Europa, en Estados Unidos no existe una norma que regule la protección de las bases de datos. Durante el siglo XIX, dos ideas fundamentales sirvieron de justificación para la protección de las bases de datos: la primera fue la conciencia del esfuerzo e inversión realizados por la persona que recopila los datos, lo que se conoce como la doctrina *sweat of the brow* (sudor de la frente); la segunda, el reconocimiento del criterio del autor en la elaboración y la selección de la información que se incluye en la base de datos.

La decisión *Feist Publications v. Rural Telephone Service Co.* marcó un cambio de criterio en Estados Unidos. Desde ese momento, se volvió necesario acreditar una cierta originalidad para reclamar los derechos titularidad. El titular podía recurrir a cuatro sistemas de protección: 1) el *copyright* —bien con una cierta originalidad en la

estructura, bien introduciendo datos que estén sujetos a *copyright*—, 2) la introducción de cláusulas contractuales, 3) la ley estatal de apropiación indebida y 4) las medidas técnicas. Una alternativa adicional, introducida más adelante, fue la ley antipiratería.

En múltiples ocasiones, Estados Unidos se ha planteado si era necesario revisar la protección de las bases de datos e igualarla con el nivel de protección que existe en Europa. Una y otra vez, la conclusión ha sido la misma: un mayor nivel de protección perjudicaría el desarrollo de la nueva economía. El exceso de protección podría frenar o dificultar el desarrollo del conocimiento. De hecho, consideran que la protección del *copyright* debe irse limitando. Así, el acceso abierto empieza a construirse como una forma alternativa.

A.4.5. Conclusiones del análisis de ambos sistemas

Del análisis de estos dos modelos, surge la cuestión de cuál de los dos resulta más adecuado. El análisis caso por caso del sistema americano parece más adecuado para la velocidad que requiere hoy el mercado. En efecto, tras el estudio realizados, la conclusión es que el sistema europeo tiene dos limitaciones fundamentales: por una parte, la complejidad de la elaboración de normas en la Unión Europea y, por otra, la fórmula continental de establecer una norma general y una lista tasada de excepciones no resulta adecuada. La Unión Europea trata de acercarse al sistema de *fair use* norteamericano mediante de la creación de las limitaciones y exenciones, donde recoge, o intenta recoger, la particularidades del *fair use*.

Más adaptabilidad y flexibilidad da el sistema de *fair use*. Dado que en Estados Unidos la existencia de monopolios no es deseable, aunque puedan reconocerse en determinados casos como es el supuesto del *copyright*, con el sistema de *fair use* introduce excepciones a ese monopolio. Por ello, permite la cita o incorporación, legal y no licenciada, de material protegido en un trabajo de otro autor, siempre que se cumplan cuatro condiciones:

- El propósito y el carácter del uso en disputa.
- La naturaleza de la obra protegida con derechos de autor.
- La importancia de la parte utilizada en relación con la obra en total.
- El efecto de dicho uso en el mercado sobre el valor de la obra protegida con derechos de autor.

Los datos de biodiversidad merecen una consideración especial puesto que, aunque incluidos en la categoría de datos científicos, tienen unas características propias. En lo que respecta a los datos, cada vez más voces defienden no poner barreras a su acceso y reutilización. En el caso de los datos de biodiversidad, aún con más razón,

la tendencia es, cada vez más, limitar su protección y fomentar que se compartan. De hecho, ya se ha defendido que, si reúnen ciertas condiciones, estos datos no son susceptibles de protección ni de apropiación.

A.4.6. La aparición de un nuevo marco de intercambio de datos: las tendencias open

Los derechos de propiedad industrial nacieron como una forma de garantizar la innovación, mediante la publicación de los datos. Hoy, sin embargo, están promoviéndose formas distintas de facilitar el acceso y la publicación de los datos.

También es importante promover que no se usen herramientas con licencias propietarias a la hora de almacenar y gestionar este tipo de datos. La protección de los datos de biodiversidad puede paralizar el intercambio de los mismos, lo que haría imposible afrontar de forma global las cuestiones que suscita: hacer frente a desastres, modelización, mediciones de pérdida de biodiversidad, etc. Aunque el problema jurídico está en revisión, muchas bases de datos de biodiversidad exigen como condición previa de acceso poner los datos en abierto, permitir su accesibilidad y reutilización.

Con esta preocupación de que son necesarias soluciones mucho más globales, empiezan a aparecer nuevas herramientas o nuevos enfoques, los conocidos con el nombre de *open* (abiertos). Aparece todo el movimiento denominado *open* en todas sus representaciones: *open access* (acceso abierto), *open data* (datos en abierto), *open science* (ciencia en abierto), *open source*, etc.

Este movimiento *open*, dejando a un lado el debate sobre la existencia de los derechos de propiedad intelectual o *copyright*, se centra en la accesibilidad y reutilización de los datos. Para ello, promueve que los autores «renuncien» a su aplicación, es decir, abran el uso de los datos lo máximo posible.

El proceso de *open access* se ha consolidado con las «tres bes» o BBB, que representan las iniciales de Budapest (Budapest Open access Initiative, 2002), Bethesda (Bethesda Statement on Open access Publishing, 2003) y Berlín (Berlin Declaration on Open access to Knowledge in the Sciences and Humanities, 2003), lugares en los que se celebraron las reuniones de las tres declaraciones que sirvieron de base para la construcción del *open access*.

La *open initiative* de Budapest (febrero de 2002) señala:

Por acceso abierto [...], entendemos su disponibilidad gratuita en Internet, para que cualquier usuario la pueda leer, descargar, copiar, distribuir o imprimir, con la posibilidad de buscar o enlazar al texto completo, recoger los artículos para su indexación, pasarlos como datos para *software* o utilizarlos para cualquier otro propósito legítimo, sin más barreras financieras, legales o

técnicas que aquellas que supongan acceder a Internet. El único límite a la reproducción y distribución de los artículos publicados y la única función del copyright en este marco, tiene que ser garantizar a los autores el control sobre la integridad de su trabajo y el derecho a ser reconocido y citado.

La Declaración de Bethesda sobre Publicación de Acceso Abierto (abril de 2003) da un paso más y exige la concurrencia de dos condiciones para que exista acceso abierto: (1) el permiso de los titulares de los derechos dando acceso libre, irrevocable, universal y perpetuo y una licencia para copiar siempre con atribución y (2) una versión completa de la obra y los materiales adicionales, en un estándar adecuado.

Finalmente, la Declaración de Berlín sobre el Acceso Abierto al Conocimiento en Ciencias y Humanidades (octubre de 2003) tiene como misión diseminar el conocimiento, mediante el uso del paradigma del acceso abierto. Por ello, cumplirá su propósito si la información se pone a disposición de la sociedad de manera expedita y amplia. Es necesario fomentar la diseminación del conocimiento utilizando, también, el paradigma de acceso abierto a través de Internet. En esta declaración el acceso abierto se define como una amplia fuente de conocimiento humano y patrimonio cultural aprobada por la comunidad científica.

Para que pueda alcanzarse la visión de una representación del conocimiento global y accesible, la web del futuro tiene que ser sostenible, interactiva y transparente. El contenido (*open data*) y las herramientas de *software* (*open source*) tienen que ser libremente accesibles y compatibles.

Sin lugar a dudas, el *open access* puede contribuir a mejorar la comunicación científica haciendo el conocimiento accesible, sin restricciones y sin coste. El concepto de *open access* se ha construido en dos fases: la digitalización del contenido, multiplicando la posibilidad de difusión del mismo, y la posibilidad de acceso sin restricciones al mismo.

Hay distintas modalidades de acceso abierto y en la actualidad se diferencian básicamente dos tipos: el acceso abierto verde (*Green open access*) —los autores directamente comparten la información poniéndola en un repositorio— y el acceso abierto dorado (*Golden open access*) —se hace accesible mediante revistas de acceso abierto revisadas por pares—. Aunque el primero es el menos costoso y el que más se está usando, el segundo es el más sostenible a largo plazo.

Los FAIR Principles se estructuran como un conjunto de principios rectores con la finalidad de hacer los datos fáciles de encontrar (*findable*), accesibles (*accessible*), interoperables (*interoperable*) y reutilizables (*reusable*). Son unos principios de alto nivel que podrán servir de marco de referencia. De hecho, muchas organizaciones están adoptando alguno de estos principios. En 2016 el G20 refrendó los FAIR Principles

y manifestó su deseo de apoyar los esfuerzos adecuados para promover la *open science* y contribuir a fomentar el acceso a la investigación financiada con fondos públicos mediante la aplicación de los FAIR Principles.

Cada uno de los principios exige, para su aplicación, una serie de condiciones clave para que se cumplan:

Box 2 | The FAIR Guiding Principles

To be Findable:
 F1. (meta)data are assigned a globally unique and persistent identifier
 F2. data are described with rich metadata (defined by R1 below)
 F3. metadata clearly and explicitly include the identifier of the data it describes
 F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
 A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 A1.1 the protocol is open, free, and universally implementable
 A1.2 the protocol allows for an authentication and authorization procedure, where necessary
 A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
 I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
 I2. (meta)data use vocabularies that follow FAIR principles
 I3. (meta)data include qualified references to other (meta)data

To be Reusable:
 R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 R1.1. (meta)data are released with a clear and accessible data usage license
 R1.2. (meta)data are associated with detailed provenance
 R1.3. (meta)data meet domain-relevant community standards

Figura 115. Cómo aplicar los principios FAIR. Fuente: Wilkinson *et al.*, 2016.

Los principios no se agotan en las cuatro palabras que los definen. Están pensados para ayudar al que tiene el dato y quiere hacerlo público. La idea es que la persona que tiene el dato y lo pone en abierto o a disposición de todos, lo comparta sin que ello signifique que asuma la obligación de mantenerlo. Recientemente, ha quedado clara la necesidad de que exista apoyo de estructuras públicas. El hecho de que los individuos apliquen los principios no es suficiente para garantizar la apertura de los datos. Es necesario que existan repositorios que alberguen los datos y que contribuyan a la aplicación de los principios. No parece que tenga sentido imponer esta carga al que tiene el dato y quiere ponerlo a disposición de todos. Permitir el uso es una cosa. Hoy ya el problema es que si se cierra la Institución o la web en la que se pone el dato, este deja de ser reutilizable.

La existencia de *FAIR* es, sin duda, el primer paso para el que tiene el dato. Sin embargo, para que se haga realidad, no es suficiente, es, además, necesario que existan infraestructuras. Con esta finalidad aparecen las *data management policies*. Esto es lo que deberían hacer los *European Open Science Cloud (EOSC)* y los repositorios

públicos, establecer unas políticas de gestión de los datos dentro de los repositorios. Un ejemplo puede encontrarse en GEOSS, en las *data management policies branding*.

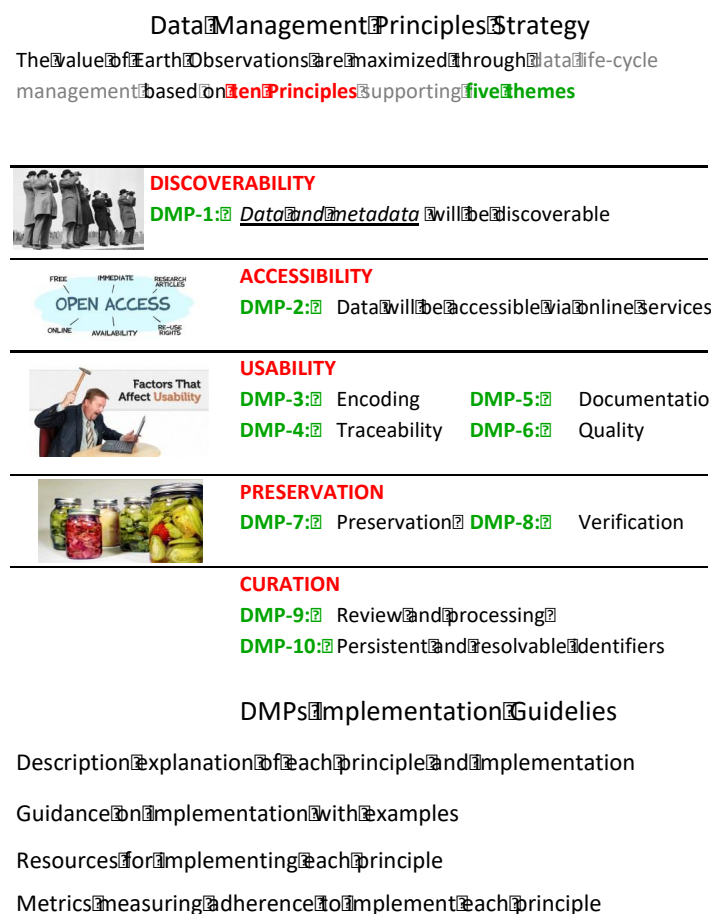


Figura 116. Data management principles branding. Fuente: Halpern, 2015.

No puede obligarse al generador del dato a reprocesarlo con las actualizaciones tecnológicas. Lo único que se le puede exigir es que lo entregue en condiciones *FAIR*. El branding permite saber cuales son los principios que se han cumplido y facilitar la reutilización de los datos al conocerse las condiciones del mismo.

Dentro de este marco, aparece el *open data*. Para que los datos estén abiertos tienen que estarlo en dos niveles: en el nivel técnico y en el nivel jurídico. Son las condiciones básicas de la existencia de interoperabilidad, que permite que los datos puedan combinarse, pues los datos de forma aislada no son nada. Por eso, el concepto se está construyendo en el proceso de definir de qué manera los datos científicos pueden y deben publicarse y reutilizarse sin coste y sin barreras. Pretende ser la base de la construcción de un concepto de ciencia mucho más globalizado.

El siguiente paso será el *open source* (código abierto), que es el *software* distribuido y desarrollado libremente. El término «libre» no solo recoge el concepto de «sin coste», sino, sobre todo, la capacidad de poder modificar la fuente del programa sin

restricciones de licencia. En efecto, muchas empresas de *software* encierran su código, ocultándolo y reservándose todos los derechos. Así, el uso de *software* cerrado puede ser una barrera para los datos. Un ejemplo podría ser el tema del GIS, cuyo uso está restringido al pago de una licencia, es decir, que los datos almacenados en ese formato solo pueden consultarlos los que dispongan de esa licencia.

El uso de programas de código abierto contribuye, también, a la diseminación de los datos y del conocimiento. El hecho de que exista libertad de acceso al código es una condición necesaria, pero no suficiente. Existen, además, toda una serie de requisitos adicionales que pueden consultarse con más detalle en el cuerpo del trabajo. Por todo lo anterior, ha quedado manifiesto que la apertura es crítica para la ciencia. Por ello, es necesario articular modelos de intercambio abierto de datos, lo que es clave en relación con los datos de biodiversidad.

Aunque el *open access* está desarrollándose a escala global, existen distintas formas de abordarlo. En este trabajo, se han contrapuesto dos regímenes diferenciados: el de Europa y el de Estados Unidos. Toda una serie de iniciativas de la Agenda Digital de Europa afectan a la construcción de este marco legislativo con la intención de ir potenciando la tendencia de poner datos en abierto. Aunque se comienza abordando el *open access* en Europa, la realidad es que el movimiento se inició un poco antes en Estados Unidos.

A.4.6.1. La construcción del *open access* en Europa

El ejemplo más representativo es el de la directiva sobre la reutilización de la información del sector público (Directiva PSI) (Parlamento Europeo y Consejo de la Unión Europea, 2003), que establece un marco general a nivel europeo con un grado mínimo de armonización. Afecta a los datos abiertos gubernamentales (mayores generadores y recabadores de datos). Los datos que gestionan, que son clave para la Administración, también pueden serlo para otros colectivos.

Se pretende, con ello, contribuir al crecimiento económico y a la creación de empleo, liberando el potencial económico de los datos, propiedad de las Administraciones Públicas, gracias a una mejora de las condiciones de explotación. El objetivo queda recogido en el artículo 1 y no es otro que establecer un conjunto mínimo de normas que regulen la reutilización y los instrumentos prácticos que faciliten la reutilización de los documentos existentes conservados por organismos del sector público de los Estados miembros. Revisada en 2013, una de las novedades principales que introduce es la recomendación de publicar la información preferentemente en formatos estándar, abiertos y procesables de modo automatizado, con el mayor nivel de granularidad posible.

Otra de las iniciativas clave es la llamada Directiva INSPIRE (Parlamento Europeo y Consejo de la Unión Europea, 2007) exclusivamente referida a mapas y datos geográficos, que exige a las autoridades públicas en toda la Unión Europea que den acceso a sus conjuntos de datos medioambientales de manera uniforme e interoperable. Una de las exigencias de la Directiva INSPIRE es la creación de un GeoPortal que permita compartir los datos y eliminar todos los obstáculos posibles. Pivota sobre cinco pilares fundamentales: los metadatos, la interoperabilidad, los servicios de redes, las políticas de datos y la implementación. De hecho, un ejemplo de los problemas que lleva aparejada la falta de interoperabilidad de estos datos es que en la línea de fijación de costa existen más de ocho diferentes. La interoperabilidad de los datos, permitiría que existiera una única representación de la línea de costa y no ocho variaciones distintas de la misma en función de donde provengan los datos. Otro ejemplo de esto, podría ser la existencia de ocho mil municipios en España, con softwares no interoperables. Escalado a toda Europa, da un indicio de los problemas que supone la no interoperabilidad técnica en la elaboración de mapas.

Para garantizar la interoperabilidad, exige la adopción de normas de implementación comunes para los metadatos, las especificaciones de datos, los servicios de red y el intercambio de datos dentro de la infraestructura paneuropea de datos espaciales.

Además, en la Unión Europea, el modelo de *open access* ha ido introduciéndose en los Programas Marco de Financiación de la Unión Europea. Aunque empezó en el 7.º Programa Marco como proyecto piloto, su consolidación se está produciendo en el marco del siguiente, el llamado Horizonte 2020.

El último de estos programas marco es Horizonte 2020, en el que ha habido un fuerte esfuerzo de financiación de proyectos realizados conjuntamente por la Comisión Europea y por el Banco Europeo de Inversiones. Horizonte 2020 marca además un cambio muy claro de política en lo que respecta a los datos abiertos. En efecto, en el nuevo marco de Horizonte 2020 implantado a partir de 2017, el acceso abierto deja de ser una opción y prevalece. Se aplica, además, de forma obligatoria y con carácter retroactivo. Adicionalmente, la Unión Europea está realizando enormes esfuerzos de financiación a través de los fondos estructurales para las nuevas infraestructuras de investigación. Obtener financiación de los programas marco tiene un impacto en los resultados y en lo que pueda o deba hacerse con ellos.

Establece el *open by default* (acceso abierto por defecto). Tres razones principales podrían amparar que no se aplicase el acceso abierto: la protección de la privacidad, la existencia de derechos de propiedad intelectual y, la tercera, más genérica, que pueda afectar al objetivo principal del proyecto. Deberá buscarse un equilibrio entre el acceso abierto y la protección total.

La Unión Europea ha elaborado una agenda digital para Europa, una de las iniciativas clave de la Estrategia 2020 para lograr un crecimiento sostenible, inclusivo y eficiente (Comisión Europea, 2010). La creación de un mercado único digital es clave, pero de momento la agenda identifica los obstáculos que dificultan la construcción de la Agenda Digital. Son los siguientes:

- La fragmentación de los mercados digitales.
- La falta de interoperabilidad.
- El incremento de la ciberdelincuencia y el riesgo de escasa confianza en la redes.
- La ausencia de inversión en las redes.
- La insuficiencia de los esfuerzos de investigación e innovación.
- Las carencias en la alfabetización y la capacitación digitales.
- La pérdida de oportunidades para afrontar los retos sociales.

A.4.6.2. Estados Unidos y las políticas de *public access*

El modelo estadounidense de acceso abierto parte de las llamadas políticas de *public access*, que es el término con el que se habla de *open access* en Estados Unidos. También en Estados Unidos, fruto de la multiplicación de datos en Internet y del encarecimiento de las suscripciones, surgen voces reclamando acceso abierto para las publicaciones científicas.

La primera referencia es la que hace el National Institute of Health en una de sus publicaciones. Establece un mandato de acceso abierto para toda investigación que subvencione con la obligación de introducirlo en *PubMed Central*, un archivo gratuito de artículos accesible a cualquiera, desde cualquier parte, mediante un simple buscador web. La totalidad del texto de los artículos es de acceso libre para su lectura, aunque pueden variar las condiciones de reutilización.

Todas las publicaciones tienen que estar obligatoriamente disponibles en el repositorio en los doce meses siguientes a su publicación. Como se puede apreciar, la importancia del acceso abierto es cada vez más fuerte en Estados Unidos.

La innovación alcanzada con fondos públicos no puede protegerse, por lo que rara vez llega a usarse con fines comerciales y, por ende, no acaba de ser productiva. También, la National Science Foundation aplica políticas que en Estados Unidos se llaman de *public access*. Como ya se vio, es fundamental que la investigación financiada con fondos públicos sea accesible.

En 2013 la Administración de Obama, recogiendo estas líneas maestras de *public access*, publica en la página web de la Casa Blanca una carta que recoge el compromiso de esta Administración con el hecho de que los ciudadanos merecen un fácil acceso a los resultados de la investigación financiada con sus impuestos (Anon., 2013).

Con el cambio de Administración y la llegada de Trump a principios de 2017, empiezan a surgir dudas sobre lo que ocurrirá con los resultados de investigación, con la propia financiación. Todo lo aprobado o publicado en la página web de la Casa Blanca con Obama ha desaparecido con Trump. Según avance la legislatura, se verá qué medidas adopta Estados Unidos en relación tanto con la apertura de los datos como con la financiación de la investigación. Aunque en un principio, parecía que se iban a producir cambios drásticos, por el momento, no parece que los haya habido. De hecho, en los últimos meses la Administración Trump ha hecho hincapié en la importancia de la investigación, de la subvención de la misma y en la necesidad de dar publicidad a los datos. Aún quedan por definir las políticas o la forma en la que esto se llevará a cabo.

En la era Trump, además, desaparece el concepto de neutralidad de la red. La red neutral⁸⁸ es aquella que permite comunicación de punto a punto sin alterar su contenido. Por lo tanto, la neutralidad de red es el principio por el cual los proveedores de servicios de Internet y los gobiernos que la regulan deben tratar a todo tráfico de datos que transita por la red de igual forma indiscriminadamente, sin cobrar a los usuarios una tarifa dependiendo del contenido, página web, plataforma o aplicación a la que accedan. Ni según el tipo de equipamiento, dispositivo o método de comunicación que utilizan para el acceso. En definitiva, la neutralidad de la red evita que el proveedor contratado cargue tarifas adicionales por visitar alguna web a su cuenta por el servicio prestado. Si este principio es vulnerado, se perderá, no solo la libertad de ver lo que quieras en la red, sino que, además, te cobrarán más.

Esta variación sobre el funcionamiento de Internet, puede amenazar la propia existencia de Internet. En efecto, uno de los pilares fundamentales sobre los que se ha ido construyendo desde su creación ha sido precisamente sobre el concepto de neutralidad. En los próximos años, se verá de que manera impacta la desaparición de la neutralidad de la red. De hecho, existen voces que piden el restablecimiento de la neutralidad derogada.

Aunque, con carácter general los datos de la investigación subvencionada no pueden protegerse, existen casos en los que puede ser deseable que el sector público explote

⁸⁸ Extraído de https://es.wikipedia.org/wiki/Neutralidad_de_red. (Consultado por última vez el 29 de marzo de 2018).

con fines comerciales los resultados obtenidos en el marco de la investigación pública.

Existen diferentes niveles de protección de los datos y de los resultados obtenidos en función del origen de la subvención. Así, en el primer nivel de protección, los datos obtenidos en proyectos subvencionados por la National Science Foundation o el National Institute of Health, por ejemplo, deberán ponerse en el dominio público, puesto que los financiados con dinero de los contribuyentes, en principio, no son susceptibles de protección. Por lo tanto, los datos y la información generada en el marco de un proyecto del Gobierno entrarán en el dominio público transcurrido un año.

En efecto, la normativa estadounidense pone unos límites a los derechos de propiedad intelectual o *copyright* y excluye su aplicación a los datos generados por órganos o agencias gubernamentales. La consecuencia es una tendencia cada vez mayor a compartir los datos abiertamente y archivarlos en repositorios públicos.

Esta norma no existe en ninguna otra parte, por lo que resulta muy difícil que Estados Unidos pueda hacer un uso similar de datos extranjeros, puesto que, en ningún otro sitio, el concepto de dominio público es tan amplio como lo es allí. A pesar de que la norma establece un libre acceso a los datos, existen factores que pueden imponer limitaciones a dicho acceso. Son fundamentalmente de tres tipos: problemas de seguridad nacional, la necesidad de proteger los datos personales de seres humanos que tomen parte en la investigación (aunque sobre este punto ya se están alzando voces para contraponer el interés de la persona al bien común) y, por último, la necesidad de respetar información de carácter confidencial (por ejemplo: por tratarse de datos originados en el ámbito del sector privado, por compañías, etc.). Acceso abierto no es sinónimo de gratuito y, así, los precios de acceso pueden variar. Sin embargo, no debe olvidarse que exigir costes superiores a los marginales puede crear sustanciales barreras de acceso a la información.

Otra cuestión a estos efectos es la disponibilidad de los datos. Con independencia de que los datos estén en el dominio público, solo si pueden encontrarse, podrán reutilizarse. Por eso, se fomenta la existencia de un repositorio desde donde los datos sean accesibles garantizando, de esta manera, la posibilidad de reutilizarlos. Por eso, la barrera fundamental a efectos prácticos que existe en Estados Unidos es que las agencias, *de facto*, ni diseminan los datos ni se ocupan de su preservación a largo plazo.

Además, el Gobierno tiene una prohibición de competir con el sector privado en la diseminación de datos. No podrá publicar aquellos a los que tenga acceso a través del sector privado, salvo que exista algún tipo de exención para hacerlo.

En un segundo nivel, los datos obtenidos de organizaciones no públicas (académicas o entidades sin ánimo de lucro), sí podrán protegerse. Será decisión de los investigadores ponerlos o no en el dominio público. Si bien tienen la capacidad de decidir si los quieren poner en el dominio público, existe una recomendación para hacer los datos lo más accesibles y reutilizables posibles, es decir, en la medida en que ello sea posible, ponerlos en el dominio público. En este caso, es solo una recomendación.

Este puede ser el caso de investigaciones realizadas en el marco de organizaciones académicas e instituciones sin ánimo de lucro que también subvencionan investigación, por ejemplo: el National Research Council (Consejo Nacional de Investigación de Estados Unidos). Los datos e información resultante de esta investigación no quedan, automáticamente y por defecto, en el dominio público. De hecho, el National Research Council recomienda que se impongan el mínimo de restricciones posibles a estos datos y a esta información y, preferiblemente, que se pongan en el dominio público cuando sea posible.

En un tercer nivel, los datos no solo son susceptibles de protegerse, sino que, con carácter general, se protegen. El marco de generación de este tipo de datos sería la colaboración entre el sector público y el sector privado, pues este último es un gran consumidor de datos del dominio público. Una normativa específica regula la relación entre el sector público y el sector privado. Aunque en este marco existen más restricciones, como se verá después, algunas razones permiten realizar excepciones, como puede ser el caso de las *march-in laws*.

El objetivo fundamental es promover la inversión privada y la innovación. No obstante, para salvaguardar al público en general del posible perjuicio que pudiese resultar del hecho que una empresa no hiciera uso o hiciera un uso no razonable de una invención financiada con fondos públicos, se permite al Gobierno estadounidense imponer dos reservas. La primera contempla que el Gobierno federal retiene una licencia no exclusiva, no transferible y gratuita para usar la patente para Estados Unidos y en su nombre. La segunda expone que la agencia federal financiadora puede, siempre que se reúnan ciertas condiciones, ejercitar sus derechos de *march-in* y obligar al titular de la patente a conceder una licencia a un solicitante responsable. Podrá ejercitar ese derecho directamente o a instancia de parte.

Las *march-in laws* se configuran como una excepción a las políticas de *public access*. En efecto, en Estados Unidos la investigación pública puede dar lugar a entrega del derecho de explotación a las compañías con la finalidad de que la investigación se desarrolle hasta el final y beneficie a los ciudadanos. Sin embargo, habida cuenta que el punto de partida son datos obtenidos con financiación pública, el Estado se reserva un derecho de expropiación sin pago (justificado porque está basado en una

concesión previa). De alguna manera, el Estado se reserva el derecho de hacerse cargo de una explotación que no está explotando correctamente la empresa a la que se le concedió. Funciona como una suerte de plazos de embargo de las patentes o de los derechos de propiedad intelectual. Nunca se han ejercitado en los Estados Unidos, y ello ha generado muchas críticas. Más allá de todas estas reflexiones, el *open* en cualquiera de sus acepciones parte del reconocimiento de la existencia de derechos de propiedad intelectual a los autores o descubridores en el ámbito de la ciencia. Sin embargo, ya hay cuestionamientos de hasta qué punto el sistema de propiedad intelectual es apropiado en la sociedad del conocimiento, donde, en realidad, el conocimiento se hace de modo colaborativo. ¿Está quedando obsoleto el sistema de propiedad intelectual o *copyright* en este nuevo paradigma? En Estados Unidos, parece que, sin haber llegado a hacer desaparecer el *copyright*, van limitando su aplicación.

A.5. El data sharing o intercambio de datos

GEOSS

Después del análisis de ambos modelos o sistemas diferenciados de abordar estas cuestiones, procede analizar todas las cuestiones de intercambio de datos o *data sharing*. Sobre este punto se analizarán distintos esquemas de intercambio de datos como puede ser el previsto en GEOSS o en CODATA enfocados principalmente en datos de investigación o científicos que son los que son objeto de esta tesis.

La ciencia es cada vez más colaborativa y más intensiva en datos. Si bien el acceso a estos permite adquirir nuevo conocimiento, sigue faltando la parte colaborativa: el *data sharing* o intercambio de datos. Compartirlos es, sin duda, el nuevo paradigma de hoy, el que permitirá un avance de la ciencia mayor que el sucedido hasta ahora. De hecho, el concepto de «colaborativo» está permeando, por ejemplo, con la economía colaborativa entre otras: la individualidad está diluyéndose.

Compartir datos es ponerlos a disposición de los demás para que puedan reutilizarlos, con el mínimo de restricciones posibles, pero este proceso no es tan sencillo como parece. Además de la reticencia de los científicos —que temen ser cuestionados, que otros se apropien de sus datos y un sinnúmero de razones más que se han tratado con detalle en el cuerpo del trabajo—, existen una serie de barreras.

Clave en este ámbito es el intercambio de datos que, en el marco de la interoperabilidad de datos científicos sobre la biodiversidad, facilitan organizaciones como el GEOSS. Una de sus preocupaciones clave es mejorar los marcos de interoperabilidad. Los principios de intercambio del GEOSS son los siguientes:

1. Se compartirán de forma total y abierta los datos, metadatos y productos, reconociendo los instrumentos internacionales relevantes y las políticas nacionales así como la regulación.
2. Todos los datos compartidos, metadatos y productos estarán disponibles en los periodos de tiempo más cortos y con el mínimo coste.
3. Todos los datos compartidos, metadatos y productos se entregarán sin cargo o con un cargo que no supere el coste de reproducción para promover la investigación y la formación.

CODATA

CODATA, una organización creada hace más de cuarenta años para promover y fomentar, a nivel mundial, la recopilación, evaluación y difusión de datos numéricos confiables de importancia para la ciencia y la tecnología, trabaja para mejorar la calidad, fiabilidad, gestión y accesibilidad de los datos relevantes en todos los ámbitos de la ciencia y de la tecnología.

CODATA proporciona a científicos e ingenieros acceso a actividades de datos internacionales para aumentar la concienciación, la cooperación directa y el desarrollo de nuevo conocimiento. Su razón de ser es contribuir al desarrollo y progreso de la ciencia y de la tecnología mediante el desarrollo y el intercambio del conocimiento en relación con los datos y con las actividades relacionadas con los mismos.

La interoperabilidad legal

La interoperabilidad legal permite conocer las condiciones jurídicas de uso de los datos. Sin duda, las discrepancias entre sistemas jurídicos pueden convertirse en barreras legales en lo que respecta a los mismos. Estas cuestiones se van a ir desgranando a medida que se aborden los distintos temas que la construyen. Con la tecnología el dato es a día de hoy un elemento fundamental; es un activo, la materia prima de la economía del conocimiento y las empresas los capturan y organizan para su explotación. Acotar su definición resulta relevante; sin embargo, lo que pudiera parecer sencillo no lo es tanto. En efecto, el problema fundamental que plantea la interoperabilidad legal es la forma en que deben gestionarse los metadatos de millones de artículos publicados. En efecto, la gestión de los metadatos jurídicos puede suponer la aplicación de condiciones jurídicas que supongan restricciones al uso de los datos, bien mediante el uso de condiciones contractuales, bien mediante restricciones aplicadas mediante el juego de los derechos de propiedad intelectual o *copyright*...La interoperabilidad legal, por lo tanto, pretene informar de las condiciones jurídicas vinculadas al uso del dato y a establecer las condiciones que

permitan mezclar datos que tengan distintas condiciones jurídicas y de qué manera se puede hacer uso de los mismos.

La interoperabilidad legal resuelve como llegar al origen. Pretender dar solución a como informar de las condiciones jurídicas del dato, de si se puede usar y de cómo puede usarse. El uso de datos con distintas condiciones jurídicas se complica. Para dar respuesta a esto nacen las licencias *Creative Commons*, que, en realidad, más que licencias *waivers*.

Las discrepancias entre sistemas jurídicos pueden convertirse en barreras legales en lo que respecta a los datos. La función de la interoperabilidad legal de datos es facilitar que los sistemas jurídicos no limiten o limiten lo mínimo posible la circulación de la investigación científica. Es la condición previa para la optimización del uso de los datos científicos, de investigación, existentes hoy que, como se vio, es uno de los componentes del *open access*, el primero de los FAIR Principles.

La clave de la interoperabilidad legal no es tanto la modificación de las normas, que tomará su tiempo, como un cambio cultura: promover que sean los propios titulares de los datos los que los pongan a disposición de la comunidad, promover las políticas de acceso abierto, como se ha visto con anterioridad, dar acceso a los datos desde el punto de vista jurídico.

La construcción de la interoperabilidad legal y de los cambios de paradigma en los que el mundo está inmerso requiere de una aproximación diferente. Por consiguiente, la interoperabilidad legal es la que trabaja en la supresión de las barreras jurídicas haciendo posible que los datos sean interoperables.

Con la finalidad de contribuir a la simplificación y mejor comprensión de esta nace todo el movimiento *copyleft*, como una forma de permitir a los autores, no especializados en cuestiones técnicas de propiedad intelectual, poner los trabajos y datos de su propiedad a disposición de todos para su reutilización. Un ejemplo que no podía dejar de mencionarse en este trabajo son las licencias Creative Commons.

Estas iniciativas crean una serie de licencias con condiciones de reutilización de los datos. Tienen una ventaja fundamental: quieren facilitar herramientas que fomenten la reutilización de los datos. Promueven, con esta finalidad, limitar al mínimo las restricciones de uso.

Aunque los datos son la base de la ciencia (recabarlos, analizarlos, publicarlos, reanalizarlos, cuestionarlos y reutilizarlos), existen barreras que no permiten acceder a ellos, que imponen restricciones de uso (de editores o suministradores de datos), y que, por ello, dificultan la reutilización de los mismos. Todavía, en los datos científicos, existe una voluntad muy marcada de reconocimiento que no ayuda.

Como respuesta a ello, una serie de iniciativas que trabajan en paralelo empiezan a elaborar principios y directrices con vistas a eliminar las restricciones jurídicas que puedan existir. Se va a proceder a analizar los más relevantes.

Después de un análisis preliminar de los principios de GEOSS, se observa que este propone la elaboración de un conjunto de directrices que regulen no solamente el intercambio de los datos, sino también su gestión. Una buena gestión de datos implica una serie de actividades que aseguran que los datos sean accesibles y puedan ser entendidos y utilizados.

Por eso, sus principios recogen cuestiones como la capacidad de encontrarlos, de acceder a ellos, su usabilidad —con referencia a sus estándares, fiabilidad o procedencia—, preservación y conservación. En primer lugar, el Foro de Belmont, integrado por un grupo internacional de agencias enfocadas a la investigación relacionada con el cambio global, es la mayor financiadora en proyectos de esta índole. Pretende acelerar la respuesta a las necesidades acuciantes en materia de investigación medioambiental suprimiendo las barreras más críticas en lo que respecta a la sostenibilidad alineando y movilizand recursos internacionales.

En segundo lugar, los Principios Panton (*Principles for Open data in Science*) parten de la asunción de que la ciencia está basada en la construcción, la reutilización y la crítica abierta de las publicaciones científicas. Aceptan alguna excepción como el que se divulgue dónde se encuentran especies en peligro de extinción o preservar la privacidad de un paciente.

En tercer lugar, aunque una buena gestión de datos no es un fin en sí mismo, es indiscutible que es clave para conducir el descubrimiento del conocimiento, la obtención de los datos y su integración. Con este razonamiento, aparecen los FAIR Principles, que se configuran como un conjunto de principios rectores con la finalidad de hacer los datos fáciles de encontrar, accesibles, interoperables y reutilizables (los datos deben ser *findable, accesible, interoperable* y *reusable*, cuyas iniciales configuran el acrónimo FAIR).

No puede dejar de hacerse referencia a los principios de interoperabilidad legal publicados en octubre de 2016 en el marco de la RDA en cuya redacción ha participado la autora de este trabajo. Estos principios son los siguientes:

- Primero. Facilitar el acceso legal y la reutilización de los datos de investigación.
- Segundo. Determinar los derechos y responsabilidades de los datos.
- Tercero. Equilibrar los intereses jurídicos.
- Cuarto. Declarar los derechos de manera transparente y clara.

- Quinto. Promover la armonización de los derechos en los datos de investigación.
- Sexto. Proporcionar la atribución y el crédito apropiados para los datos de la investigación.

Una vez formulados, la idea es promover su adopción por parte de las organizaciones, instituciones y universidades, y es adonde se dedicarán los próximos esfuerzos.

El dominio público el movimiento copyleft y las licencias Creative Commons: primeros pasos hacia la interoperabilidad global

Por último, está la importancia del ámbito del *public domain* y lo que ello supone. El análisis realizado en este trabajo lleva a analizar cómo el hecho de poner los datos y los descubrimientos en el dominio público puede ser uno de los primeros pasos hacia la interoperabilidad legal. Con los matices que ello tiene. En efecto, una de las cuestiones que se plantea es si se puede poner en el dominio público algo sin que el titular del dominio público lo acepte. Eso supondría que nadie puede poner los datos en el Domicio público sin que los Estados, titulares del mismo, lo acepten. Un concepto que es muy claro cuando se habla de bienes inmuebles, no puede “darse al Estado” una propiedad sin que este la acepte, es más difícil de ver en el ámbito digital. Para ello, se crean *waivers* cuya finalidad es renunciar a los derechos de propiedad (intelectual/industrial) sobre unos datos y permitir que los usen terceros sin que ello necesariamente signifique que estén en el dominio público.

Esta confluencia de regímenes genera una confusión en los científicos que provoca, bien que no reutilicen los datos, bien que los usen sin respetar las normas que los protegen, como mínimo, por desconocimiento.

El acceso público a los datos se considera ya parte de la preservación de la ciencia. El primer paso para avanzar a modelos de innovación es reconocer que el dominio público representa espacios públicos que conducen a la eficiente difusión del conocimiento.

Ya se ha arrancado la construcción de la interoperabilidad con esta serie de herramientas e instrumentos analizados. Tal vez, el segundo, el verdadero cambio cultural, suponga encontrar un nuevo marco jurídico al margen de los derechos de propiedad intelectual e industrial que permita florecer esta nueva forma de construir el conocimiento de forma colaborativa.

Sin duda, quedan muchas cuestiones pendientes de abordar en la construcción de este segundo paso, en el que se diluye o, quizás, empieza a diluirse la protección.

Conocer los niveles de protección no es suficiente hoy, pues la reutilización de los datos va demasiado rápida para seguirle la pista.

¿Considerarán las nuevas generaciones que se les reconozca de forma individual lo que han conseguido? ¿Seguirá siendo el modelo de protección adecuado para lo que viene? Tal vez quede por dar el salto, el paso real que, integrando el nuevo paradigma en toda su dimensión, sea capaz de asumir que el marco de los derechos de propiedad intelectual e industrial, tal como los conocemos hoy, ya no es el adecuado. Seguir buscando excepciones y fórmulas para, manteniendo el mismo sistema, encontrar soluciones diferentes basadas en los mismos principios, tal vez no sea más un parche. Quizás sea, por tanto, necesario un nuevo marco jurídico mucho más flexible, que ya está ocurriendo en muchos ámbitos. ¿Estarán naciendo los derechos de propiedad intelectual e industrial o *copyright by design*?

B. Segundo marco contextual: la biodiversidad

B.1. El concepto de biodiversidad

Este punto arranca de la construcción histórica del concepto de biodiversidad a mediados del siglo XX. El hombre y la biodiversidad están indefectiblemente unidos, pues la supervivencia misma del hombre está unida a la protección de los distintos ecosistemas y de los servicios de los mismos.

Sin embargo, se da un punto de inflexión cuando el hombre empieza a tomar conciencia de que la naturaleza es un recurso finito. El hecho de que la población de la Tierra se haya multiplicado por seis en doscientos años es determinante (alcanza un total de seis mil millones en 2016) (Población mundial, s. f.). Las previsiones son de un crecimiento exponencial acelerado con una reducción de la mortalidad.

Como no podía ser de otra manera, el consumo de todos los recursos naturales también se ha multiplicado por seis en los últimos doscientos años. Por eso, en los años setenta, empiezan a surgir voces de alarma, primero, alertando sobre la problemática del cambio climático y, después, poniendo sobre la mesa todas las cuestiones de diversidad biológica. Surge, también, un movimiento que muestra ya una preocupación por la conservación de la diversidad biológica. Aparecen entonces algunas iniciativas intergubernamentales, por ejemplo: la Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres (CITES).

En un principio, la diversidad biológica hacía referencia a la diversidad de especies como característica estructural de los ecosistemas. En los años ochenta aparece ya el término de «biodiversidad», en el marco de una reunión auspiciada por la National Science Foundation: el National Forum on BioDiversity.

La aparición de este término lleva a repensar la definición de diversidad biológica o biodiversidad, que contempla también la diversidad genética (la cantidad de variabilidad genética dentro de las especies) y la diversidad ecológica (el número de especies en una comunidad de organismos).

La diversidad biológica, entendida como la variabilidad de seres vivos y de los ecosistemas en los que viven, es el pilar sobre el que se han construido las civilizaciones humanas, y su conservación es la base del desarrollo sostenible. Así, se pone de relieve la relación que existe entre la biodiversidad y el desarrollo económico.

Como consecuencia de toda esta construcción y análisis, los científicos se dan cuenta de los problemas que está suponiendo hoy, y puede suponer a la larga, la pérdida de biodiversidad. Todo ello, conduce a una reflexión sobre cómo abordar esta cuestión de manera global.

B.2. El CDB

Sin lugar a dudas, el Panel de Cambio Climático (IPCC) marcó un punto de inflexión. Aunque no está directamente relacionado con la biodiversidad, fue el primer paso y el modelo que, una y otra vez, la biodiversidad toma como referencia.

Abordar el problema de la pérdida de biodiversidad a escala global requirió un convenio: el CDB. En 1992 se iniciaron en el marco de las Naciones Unidas dos movimientos en paralelo: uno enfocado en el cambio climático que terminó en la Convención Marco de Naciones Unidas de Cambio Climático (aprobada en 1992 y ratificada en 1994) y otro dirigido a la diversidad biológica o biodiversidad, del que nacería el Convenio de Diversidad Biológica.

El problema en el marco del cambio climático fue solucionado desde el principio con la creación del IPCC que permitía gestionar la interrelación entre la ciencia y la toma de decisiones en materia de cambio climático. El foco y la definición clara de los problemas en materia de cambio climático ha hecho posible la creación de un modelo claro. Sin duda, las cuestiones que se plantean en el ámbito del cambio climático y su análisis es más lineal, siendo el abordarlos más sencillo. La complejidad de la pérdida de biodiversidad es evidente, tanto por todos los factores que inciden cuanto por las dificultades que existen a la hora de medir la pérdida.

El origen de las negociaciones se remonta a 1987. En la Decisión 14/86 (Naciones Unidas, s. f.) del Consejo de Administración del PNUMA (Programa de las Naciones Unidas para el Medio Ambiente) se decidió constituir un grupo de expertos en biodiversidad con el objetivo de armonizar todos los convenios existentes relacionados con el tema. Con ello, quedaba patente la necesidad de crear

un instrumento internacional vinculante para la conservación y uso sostenible de la diversidad biológica.

La firma del CDB fue histórica, pues se creó una estructura internacional para promover la cooperación internacional y la implementación en los países firmantes. Nació como un tratado internacional jurídicamente vinculante que establecía como objetivos fundamentales: la conservación de la biodiversidad, el uso sostenible de sus componentes y la participación justa y equitativa en los beneficios que se derivasen de la utilización de los recursos genéticos.

Al principio, la biodiversidad ponía el énfasis en la diversidad de especies (procedente de Darwin). Ahora se ha ampliado. De hecho, el CDB dice que la biodiversidad va de la diversidad de especies a la diversidad de ecosistemas. Queda definido en el artículo 2 del CDB que define la diversidad biológica y dice que se entiende por diversidad biológica la variabilidad de organismos vivos de cualquier fuente, incluidos, entre otras cosas, los ecosistemas terrestres y marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte: comprende la diversidad dentro de cada especie, entre las especies y los ecosistemas.

Por fin, los paisajes están empezando a considerarse incluidos también en el concepto de biodiversidad, después de muchos años de discrepancias en los que muchos biólogos defendían que los paisajes eran para los pintores. Sin embargo, ya los científicos empiezan a reconocer el paisaje como una escala superior de biodiversidad que no deja de ser la escala que agrupa los ecosistemas. Se pueden tener ecosistemas parecidos y que los paisajes sean diferentes. Cambia la estructura de donde está uno y donde está uno y donde está otro. En realidad, el paisaje influye en los ecosistemas (Tucker, M. A. *et al.* 2018). Por lo tanto, la conservación de la biodiversidad requiere también la protección de los paisajes ecológicos. En España, de hecho, la Ley 42/2007, de 13 de diciembre, del Patrimonio Natural y de la Biodiversidad ya incluye las referencias de los paisajes.

Una de las cuestiones que afloraron durante la negociación del convenio fue que, en materia de biodiversidad, la riqueza, se encuentra en los países del sur. Los países del norte, a cambio, disponen de los recursos económicos para investigar y extraer conclusiones del análisis de la biodiversidad. Por eso, en dicho convenio subyace el principio de participación justa y equitativa en los beneficios.

Forman parte del CDB más de ciento noventa Estados, aunque Estados Unidos no se ha adherido. En efecto, implicado en la creación y desarrollo técnico del convenio, se negó después a ratificarlo y sigue sin hacerlo. Dicha falta de ratificación no es falta de concienciación en materia de diversidad biológica, pues, de hecho, Estados Unidos fue un temprano impulsor de políticas de biodiversidad en el ámbito nacional y es un convencido de la preservación de la diversidad biológica.

La negativa de Estados Unidos a ratificar el CDB tiene una justificación múltiple. El entonces presidente Bush señaló como claves las cuestiones ligadas a los derechos de propiedad industrial frente al interés de las compañías norteamericanas de maximizar los beneficios en biotecnología, así como los términos previstos de ayuda a los países del sur o países en desarrollo.

Los estadounidenses, preocupados por sus intereses económicos y por el control sobre sus tierras, se siguen negando a ratificar el CDB, pues no quieren permitir una intervención externa que limite su organización interna. La fuerza de la industria estadounidense enfocada en cuestiones de patentes y los agricultores han hecho que Estados Unidos siga sin ratificar el CDB.

En el CDB existen tres órganos: el órgano rector, la Conferencia de las Partes, integrado por los representantes de los Estados y las organizaciones parte del CDB, la Secretaría Ejecutiva y el Subsidiary Body on Scientific, Technological and Technical Advice (SBBSTA). Este último nace con la finalidad de ser un órgano científico asesor.

Muy pronto se politiza, es decir, que no está integrado por científicos. Se habla incluso de PreCoP, el SBBSTA no acaba de dar los *inputs* que el CDB necesita para funcionar. Sin embargo, el CDB, más que un órgano de acción, ha resultado ser un lugar de consenso. De hecho, «delega» las medidas a las partes en el ámbito nacional y no establece objetivos concretos, sino principios generales para que las partes tengan en consideración. No se ha creado el marco científico necesario.

Es un hecho que el CDB no ha acabado de conseguir su objetivo de frenar la pérdida de biodiversidad. Se ha hablado de inadecuada implementación, de exceso de ambición e, incluso, se ha imputado el fracaso a la inexistencia de un órgano de contenido científico fuerte, fruto de la politización del SBSSTA. No obstante, la falta de consenso y la inexistencia de unidades de medida para calcular son las razones fundamentales. Desde luego, el entramado del CDB no ha ayudado a que cumpla sus objetivos.

En el marco de los CDB y para la medición de la pérdida de biodiversidad se elaboran los Aichi Targets, con la idea de tener unos indicadores de referencia para medir la pérdida de biodiversidad. Su principal reto estratégico es ser capaz de definir unos objetivos que contribuyan a frenar la pérdida de biodiversidad, teniendo en cuenta la complejidad y variedad de factores que afectan tanto en materia de biodiversidad como de ecosistemas. Si bien la idea con la que se elaboraron tenía sentido, lo cierto es que con el paso del tiempo han resultado ineficientes para conseguir frenar la pérdida de biodiversidad. La ambigüedad en la redacción de los objetivos, la complejidad y las redundancias de los mismos, así

como el hecho de no tener indicadores cuantificables asociados, ha hecho que no hayan dado los efectos esperados.

Es un hecho que el CDB no ha acabado de conseguir su objetivo de frenar la pérdida de biodiversidad. Se ha hablado de inadecuada implementación, de exceso de ambición e, incluso, se ha imputado el fracaso a la inexistencia de un órgano de contenido científico fuerte, fruto de la politización del SBSSTA. No obstante, la falta de consenso y la inexistencia de unidades de medida para calcular son las razones fundamentales.

En el marco del CDB y para la medición de la pérdida de biodiversidad se elaboran los Aichi Targets, con la idea de tener unos indicadores de referencia para medir la pérdida de biodiversidad. Su principal reto estratégico es ser capaz de definir unos objetivos que contribuyan a frenar la pérdida de biodiversidad, teniendo en cuenta la complejidad y variedad de factores que afectan tanto en materia de biodiversidad como de ecosistemas. Si bien la idea con la que se elaboraron tenía sentido, lo cierto es que con el paso del tiempo han resultado ineficientes para conseguir frenar la pérdida de biodiversidad. La ambigüedad en la redacción de los objetivos, la complejidad y las redundancias de los mismos, así como el hecho de no tener indicadores cuantificables asociados, ha hecho que no hayan dado los efectos esperados.

En efecto, resulta muy complicado lograr un esfuerzo global para el asesoramiento sobre el estatus de la biodiversidad, sin indicadores que permitan medir de alguna manera la pérdida de biodiversidad y los efectos de la misma. La necesidad de replicar el modelo de Cambio Climático en el marco de la biodiversidad sigue presente. Aunque no puede perderse de vista la gran complejidad que tiene el tema de la biodiversidad frente a las cuestiones de Cambio Climático.

Cualquier solución tiene que tener en cuenta la soberanía y las aportaciones de los locales. Los acuerdos internacionales, como el CDB, serán ineficientes salvo que consigan generar los incentivos económicos necesarios para que tanto la gente como las organizaciones y los Gobiernos se comprometan y actúen.

Fruto de que el CDB no alcanzó sus objetivos de frenar la pérdida de biodiversidad, un nuevo instrumento recoge el reto: la Plataforma Intergubernamental sobre Biodiversidad y Servicios de los Ecosistemas (IPBES, por las siglas en inglés de Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services).

El IPBES, igual que el CDB antes, pretende replicar el IPCC, sin lugar a dudas el instrumento que mejor ha funcionado y que más compromiso ha alcanzado como instrumento internacional, hasta tal punto que recibe el nombre de IPCC de la biodiversidad. Para lograr lo que el CDB no consiguió, el IPBES necesita afrontar

una serie de retos. Se configura como un instrumento de ámbito general que, al igual que el CDB, afecta a la biodiversidad y a los servicios de los ecosistemas con carácter general.

B.3. El IPBES

Como plataforma que es, el IPBES demuestra mucha más flexibilidad a la hora de abordar las cuestiones de biodiversidad. Además, su configuración permite trabajar en colaboración con muchos otros instrumentos. De hecho, es cuestión de tiempo que el IPBES sea alimentado por todos los datos de las distintas herramientas analizadas en el punto siguiente. No tiene plataforma informática. Sigue los principios clásicos.

Nació, formalmente, en 2012, como un órgano intergubernamental abierto a todos los miembros de Naciones Unidas con la finalidad de «fortalecer la interfaz entre lo científico y lo político en lo que respecta a la biodiversidad y a los servicios de los ecosistemas para la conservación y uso sostenible de la biodiversidad, bienestar de los hombres a largo plazo y para el desarrollo sostenible».

Las interfaces científico-políticas tienen que reunir cuatro elementos para poder funcionar: la investigación, la observación, la evaluación y la gobernanza. A results del fracaso del CDB, se percibe la necesidad de dar énfasis a la relevancia política, sin por ello perderlo en la evaluación y la gobernanza. Tiene un enfoque muy potente en la búsqueda de la transparencia en relación con la gobernanza de la biodiversidad.

Se aprecia, también, la falta de información que existe en materia de biodiversidad, por eso el IPBES quiere convertirse en un lugar que centralice la información obtenida de distintas fuentes. De hecho, como se verá después, el GBIF y GEO BON se ponen a su disposición, como fuentes de información. Ya se está trabajando con el IPBES para que tome conciencia de lo importantes que son estas fuentes de información para su funcionamiento. Sin embargo, el IPBES todavía está funcionando con medios tradicionales y no digitalizados y no aprovechando todas las fuentes de información que podría tener.

Sin embargo, si el marco general estaba muy claro, en el siguiente paso de definición de los marcos conceptuales, empiezan a aparecer algunas divergencias. Esta plataforma puede resultar clave, aunque el proceso de integración no está acabado, más bien al contrario, queda mucho por hacer como ha quedado puesto de manifiesto en el párrafo anterior.

LA APLICACIÓN DE LAS NUEVAS TECNOLOGÍAS DIGITALES A LA BIODIVERSIDAD: EL NACIMIENTO DE LA INFORMÁTICA APLICADA A LA BIODIVERSIDAD O BIOINFORMÁTICA

A. El nacimiento del término bioinformática

Después de esta doble contextualización (nuevas tecnologías y biodiversidad), procede analizar el impacto de la aplicación de las nuevas tecnologías a la conservación de la biodiversidad. Nace así el concepto de informática de la biodiversidad o bioinformática. Merecen una reflexión específica los orígenes y construcción de esta disciplina. Esta materia novedosa analiza la forma de aplicar todos los nuevos paradigmas de gestión de datos científicos y de investigación a los datos de biodiversidad.

El concepto, que nació como la aplicación de la informática a la biodiversidad, ha ido expandiéndose y aplicándose a otras disciplinas. La multiplicación de los datos hace necesario recurrir a las herramientas informáticas para poder analizarlos, tratarlos.

La bioinformática es un término amplio que incluye la aplicación de la tecnología (informática) a todas las ciencias. Es una rama de la informática que trata sobre la información de organismos vivos. Sus objetivos son fundamentalmente: (1) movilizar el conocimiento actual, (2) compartir dicho conocimiento y la experiencia del sinfín de implantaciones a nivel global, (3) evitar la creación de silos y que se reinventen las herramientas de implantación del conocimiento, (4) afrontar los retos adecuadamente y (5) buscar soluciones a retos difíciles de índole estratégica.

El eje sobre el que pivota es el concepto de dato asociado a tres cuestiones principales: la recogida o captura de datos, la compilación y servicios de estos y su visualización. El reto fundamental al que se enfrenta es la integración de la información digital procedente de una variedad de fuentes, condición necesaria para que proliferen la ciencia intensiva en datos.

B. Primeros pasos de la bioinformática: la creación del GBIF

Los primeros pasos de una herramienta de bioinformática pueden encontrarse en el GBIF. Nació con un tratado internacional específico sobre esta materia. Se configuró como una de las primeras herramientas en la construcción de la bioinformática y es una organización intergubernamental estructurada en nodos. Su origen puede encontrarse en una recomendación del Subgrupo de Informática de la Biodiversidad del MegaScience Forum, de 1999. Pone de manifiesto la necesidad de crear un mecanismo internacional para hacer accesibles los datos y la información

relacionados con la biodiversidad y, también, quería promover el intercambio de información digital. Primero se dirigió, sobre todo, a las colecciones de herbarios y museos de los países en desarrollo participantes. Su objetivo es facilitar el acceso —vía Internet, de manera libre y gratuita— a datos de biodiversidad de todo el mundo para apoyar la investigación científica, fomentar la conservación biológica y favorecer el desarrollo sostenible. Es una herramienta esencial para el IPBES. Es mucho más importante de lo que puede parecer en un primer momento.

Ahora dispone ya de datos de observación, con más de trescientos millones de registros primarios de biodiversidad, a los que es posible acceder a través de su portal web, y tiene un gran potencial de crecimiento. Es un proceso que trabaja con repositorios en librerías, no tiene la capacidad de hacer procesamientos a gran escala de imágenes. Tampoco dispone de laboratorios virtuales.

Se pretende dar acceso con un portal único. Alimentan la infraestructura muchas instituciones de todo el mundo. La red del GBIF incluye un registro central de nodos de datos y dispone de una red internacional de nodos, pero necesita flexibilidad para gestionar las necesidades de los proveedores de datos y de los usuarios que quieren acceder a los datos de biodiversidad.

Para lograrlo, la Secretaría del GBIF ha desarrollado una arquitectura de información. Una de sus prioridades es la estandarización, condición necesaria para lograr que, con un acceso único, puedan hacerse uso de los datos.

De hecho, el GBIF se pone al servicio del IPBES poniendo a su disposición los hallazgos, la movilización —que puede realizar gracias a la red de nodos nacionales— y el acceso a datos. Además, trabaja en la estandarización de estos con la finalidad de que sean interoperables y aplica principios de acceso abiertos. Permite, asimismo, que los datos de biodiversidad se usen para la investigación científica. Además, está activamente involucrado en otras iniciativas vinculadas a la biodiversidad como puede ser el GEO BON. Disponer de datos de biodiversidad es clave, pero sigue faltando dar un paso más y generar los indicadores para medir la pérdida de biodiversidad.

B.1. El GBIO

Otro instrumento esencial en la gestión de datos, singularmente de datos satelitales es el Sistema de Observación de la Tierra (GEO, por las siglas en inglés de Global Earth Observation). Nació como un grupo *ad hoc*, intergubernamental y copresidido por la Comisión Europea, Japón, Sudáfrica y Estados Unidos, con la misión de elaborar un plan a diez años (presentación en febrero de 2005) con el objeto de coordinar la observación de la Tierra. Con la voluntad de promover dicha observación y la captura de datos a gran escala, dentro del marco de GEO, nace el

Sistema de Sistemas de Observación de la Tierra (GEOSS, por las siglas en inglés de Global Earth Observation System of Systems). Esta organización habla de lo vivo.

Este sistema de sistemas tiene una función integradora que persigue dar coherencia a todas las recopilaciones de datos existentes en este momento. Su finalidad principal es la interconexión de sistemas. Se define como una forma de dar coherencia a la observación y de registrar valores. Uno de los posibles efectos será evitar, a futuro, que fenómenos «observados» tengan efectos devastadores para el hombre por la falta de canales de información adecuados, como ocurrió en el tsunami de Tailandia, en el que murieron doscientas sesenta mil personas.

La observación de la Tierra no es suficiente y es, además, necesario organizar toda la gestión de los datos obtenidos. Ligado a esto, el GEO identifica nueve áreas de bienestar que requieren de observación. Una de las áreas es la biodiversidad. Nace así, como iniciativa en 2008, el Grupo de Observación de la Tierra de Observación de la Biodiversidad (GEOBON, por las siglas en inglés de Group on Earth Observations on Biodiversity Observation), específicamente, el instrumento de GEO cuya misión es la observación de la biodiversidad. Se ha desarrollado como una organización internacional de reconocido prestigio y hoy cuenta con más de cuatrocientos cincuenta miembros y asociados. Tanto el CDB como el IPBES lo reconocen como una organización clave a la hora de desarrollar el seguimiento de la biodiversidad a nivel global.

La Secretaría de Operaciones de GEO BON proporciona la estructura de base de la organización. Es la responsable de fijar las reuniones y de mantener el sitio web. Persigue mejorar la disponibilidad de los datos de cambio de la biodiversidad a los que adoptan las decisiones y a los científicos con la finalidad de que puedan adoptarse políticas.

Inicia y coordina esfuerzos con la finalidad de diseñar e implementar programas de supervisión de la biodiversidad interoperables a nivel nacional y regional, y persigue construir una red de sistemas de observación que proporcione una información de biodiversidad mejorada armonizada con la finalidad de permitir una mejor toma de decisiones tanto a nivel local como global. Tiene como misión suministrar un marco global y científicamente sólido en lo que respecta a las observaciones a la hora de identificar los cambios en la biodiversidad. Asimismo, coordina la recogida y comunicación de los cambios que se produzcan a escala global en relación con el cambio de la biodiversidad.

Se compromete también a garantizar la continuidad a largo plazo del suministro de datos y a proporcionar una serie de productos innovadores y relevantes basados en la integración de conjuntos de datos clave. Una de las claves, ya analizada varias veces, a la hora de abordar la pérdida de la biodiversidad es la creación de

indicadores que permitan medir y evaluar lo que está ocurriendo. Como ya se vio, una de las razones del fracaso del CDB fue precisamente no haber sido capaz de elaborar indicadores de la pérdida de biodiversidad.

Esa necesidad, no abordada en el IPBES, es la tarea fundamental que tiene que abordar el GEO BON. Por eso empiezan a trabajar en la adopción e implementación de las variables esenciales de biodiversidad.

Como se verá después, se retoma su desarrollo en el marco de un proyecto financiado por la Unión Europea. Al igual que ocurría con el GBIF, los resultados del GEO BON servirán de base para proporcionar información a otras infraestructuras y organismos. El IPBES se alimenta también de la información que el GEO BON genera.

Todas estas organizaciones son diferentes, dado que surgieron para dar respuesta a necesidades diferentes. Mientras el CDB está dentro del marco de las organizaciones políticas, el IPBES representa, de alguna manera, la parte de asesoramiento científico, la que el CDB no pudo o no supo desarrollar en el marco del SBSTTA, que era el responsable de jugar ese papel de asesoramiento científico.

El GEO y sus derivadas recogen observaciones de la tierra en GEOSS en nueve áreas de bienestar: GEO BON está centrado en temas de biodiversidad; GBIF trabaja con datos para ponerlos al alcance de todos los científicos. El GEO BON y el GBIF son instrumentos que permiten a otras organizaciones tomar decisiones informadas; de hecho, alimentan el IPBES.

A fin de cuentas, la gestión de la gobernanza de la biodiversidad depende de la existencia de información fiable. Las distintas iniciativas tienen que ir dando coherencia a todo el proceso y, sin perder de vista la necesidad de depurar, algunas, de hecho muchas de ellas, están interrelacionadas y son interdependientes. Puede incluso decirse que son complementarias y generan sinergias dado que se mueven en marcos diferentes.

Esto no quiere decir que a medio o largo plazo puedan desaparecer algunas de ellas o integrarse dentro de otras. El dibujo está en pleno desarrollo y es totalmente dinámico.

C. GEO BON

Otro instrumento esencial en la gestión de datos, singularmente de datos satelitales es el Sistema de Observación de la Tierra (GEO, por las siglas en inglés de Global Earth Observation). Nació como un grupo *ad hoc*, intergubernamental y copresidido por la Comisión Europea, Japón, Sudáfrica y Estados Unidos, con la misión de elaborar un plan a diez años (presentación en febrero de 2005) con el objeto de

coordinar la observación de la Tierra. Con la voluntad de promover dicha observación y la captura de datos a gran escala, dentro del marco de GEO, nace el Sistema de Sistemas de Observación de la Tierra (GEOSS, por las siglas en inglés de Global Earth Observation System of Systems). Esta organización habla de lo vivo.

Este sistema de sistemas tiene una función integradora que persigue dar coherencia a todas las recopilaciones de datos existentes en este momento. Su finalidad principal es la interconexión de sistemas. Se define como una forma de dar coherencia a la observación y de registrar valores. Uno de los posibles efectos será evitar, a futuro, que fenómenos «observados» tengan efectos devastadores para el hombre por la falta de canales de información adecuados, como ocurrió en el tsunami de Tailandia, en el que murieron doscientas sesenta mil personas.

La observación de la Tierra no es suficiente y es, además, necesario organizar toda la gestión de los datos obtenidos. Ligado a esto, el GEO identifica nueve áreas de bienestar que requieren de observación. Una de las áreas es la biodiversidad. Nace así, como iniciativa en 2008, el Grupo de Observación de la Tierra de Observación de la Biodiversidad (GEOBON, por las siglas en inglés de Group on Earth Observations on Biodiversity Observation), específicamente, el instrumento de GEO cuya misión es la observación de la biodiversidad. Se ha desarrollado como una organización internacional de reconocido prestigio y hoy cuenta con más de cuatrocientos cincuenta miembros y asociados. Tanto el CDB como el IPBES lo reconocen como una organización clave a la hora de desarrollar el seguimiento de la biodiversidad a nivel global.

La Secretaría de Operaciones de GEO BON proporciona la estructura de base de la organización. Es la responsable de fijar las reuniones y de mantener el sitio web. Persigue mejorar la disponibilidad de los datos de cambio de la biodiversidad a los que adoptan las decisiones y a los científicos con la finalidad de que puedan adoptarse políticas.

Inicia y coordina esfuerzos con la finalidad de diseñar e implementar programas de supervisión de la biodiversidad interoperables a nivel nacional y regional, y persigue construir una red de sistemas de observación que proporcione una información de biodiversidad mejorada armonizada con la finalidad de permitir una mejor toma de decisiones tanto a nivel local como global. Tiene como misión suministrar un marco global y científicamente sólido en lo que respecta a las observaciones a la hora de identificar los cambios en la biodiversidad. Asimismo, coordina la recogida y comunicación de los cambios que se produzcan a escala global en relación con el cambio de la biodiversidad.

Se compromete también a garantizar la continuidad a largo plazo del suministro de datos y a proporcionar una serie de productos innovadores y relevantes basados en

la integración de conjuntos de datos clave. Una de las claves, ya analizada varias veces, a la hora de abordar la pérdida de la biodiversidad es la creación de indicadores que permitan medir y evaluar lo que está ocurriendo. Como ya se vio, una de las razones del fracaso del CDB fue precisamente no haber sido capaz de elaborar indicadores de la pérdida de biodiversidad.

Esa necesidad, no abordada en el IPBES, es la tarea fundamental que tiene que abordar el GEO BON. Por eso empiezan a trabajar en la adopción e implementación de las variables esenciales de biodiversidad.

Como se verá después, se retoma su desarrollo en el marco de un proyecto financiado por la Unión Europea. Al igual que ocurría con el GBIF, los resultados del GEO BON servirán de base para proporcionar información a otras infraestructuras y organismos. El IPBES se alimenta también de la información que el GEO BON genera.

Todas estas organizaciones son diferentes, dado que surgieron para dar respuesta a necesidades diferentes. Mientras el CDB está dentro del marco de las organizaciones políticas, el IPBES representa, de alguna manera, la parte de asesoramiento científico, la que el CDB no pudo o no supo desarrollar en el marco del SBSTTA, que era el responsable de jugar ese papel de asesoramiento científico.

El GEO y sus derivadas recogen observaciones de la tierra en GEOSS en nueve áreas de bienestar: GEO BON está centrado en temas de biodiversidad; GBIF trabaja con datos para ponerlos al alcance de todos los científicos. El GEO BON y el GBIF son instrumentos que permiten a otras organizaciones tomar decisiones informadas; de hecho, alimentan el IPBES.

A fin de cuentas, la gestión de la gobernanza de la biodiversidad depende de la existencia de información fiable. Las distintas iniciativas tienen que ir dando coherencia a todo el proceso y, sin perder de vista la necesidad de depurar, algunas, de hecho muchas de ellas, están interrelacionadas y son interdependientes. Puede incluso decirse que son complementarias y generan sinergias dado que se mueven en marcos diferentes.

EUBON

Esto no quiere decir que a medio o largo plazo puedan desaparecer algunas de ellas o integrarse dentro de otras. El dibujo está en pleno desarrollo y es totalmente dinámico.

D. Aplicación de una escala inferior a la global: dos modelos continentales de gestión de la bioinformática

D.1. El modelo europeo: LifeWatch

A continuación, se analizará cómo ha cristalizado en la e-ciencia el concepto de infraestructura, esta vez, dedicada a cuestiones de biodiversidad. Para ello, se recurrirá a dos ejemplos continentales que, a su vez, permitirán comparar dos modelos de regulación: el europeo y el estadounidense. Se analizará primero el modelo europeo.

Existe un ejemplo concreto de aplicación de la normativa europea a una infraestructura de biodiversidad: LifeWatch. Aterrizar la aplicación en un modelo concreto permitirá evaluar la gestión de la e-ciencia en Europa. LifeWatch ERIC arrancó en 2011, con la adopción de la Decisión de Ejecución (UE) 2017/499 de la Comisión, de 17 de marzo de 2017, relativa a la creación de la Infraestructura Virtual Europea de Ciencia y Tecnología para la Investigación sobre la Biodiversidad y los Ecosistemas-Consortio de Infraestructuras de Investigación Europeas (LifeWatch ERIC), que acabó la fase preparatoria.

Por lo tanto, la fecha de creación definitiva del ERIC es el 17 de marzo de 2017. La asamblea constitutiva tuvo lugar los días 8 y 9 de mayo de 2017. Pretende ser una e-infraestructura de referencia mundial para la protección, gestión y uso sostenible de la biodiversidad. Construirá la Infraestructura de e-Ciencia y Tecnología de Investigación para Datos y Observatorios de la Biodiversidad. Involucra a ocho países europeos: Italia, Grecia, Bélgica, Rumanía, Portugal, España, Eslovenia y Holanda. La primera asamblea operativa ha tenido lugar el 19 de enero de 2018. Se ha aprobado el plan estratégico. La constitución de un ERIC requiere tres miembros de los cuales la mayoría tienen que ser de Estados miembros de la Unión Europea, aunque no los tres.

En España (Sevilla) se encuentra fijada la sede estatutaria del ERIC, donde este se aloja. En Málaga, se encuentra también la coordinación y gestión de las operaciones así como las infraestructuras tecnológicas, es decir los superordenadores. Para ello, consta de superlaboratorios. La sede se encuentra en Andalucía porque el dinero procede de la Junta de Andalucía.

En Italia se fija el centro de servicios de LifeWatch. Por último, Holanda albergará los laboratorios virtuales aunque sólo a nivel de servicio de usuarios y de middleware, el centro de superordenadores y servidores está en Málaga.

Pretende dar soporte a una mejor comprensión de la biodiversidad. La arquitectura de la e-infraestructura se hará con laboratorios virtuales con la tecnología más avanzada para capturar, estandarizar, integrar, analizar y construir modelizaciones de

biodiversidad, y considerar escenarios de cambio. Reunirá: una red de observatorios marinos, terrestres y de aguas dulces; acceso único y abierto a una gran cantidad de información distribuida en bases de datos interconectadas y sitios de monitorización; instalaciones de computación en laboratorios virtuales con herramientas analíticas y de modelización y apoyo y formación a usuarios identificados y un programa de servicios públicos.

Además, en su condición de e-infraestructura persigue contribuir a las iniciativas internacionales en relación con la conservación de la biodiversidad y la gestión de la misma.

E. Dos ejemplos de Estados unidos

E.1. Data ONE

Organización de organizaciones, subvencionada por la National Science Foundation, trabaja para asegurar la preservación, acceso, utilización y reutilización de datos científicos multidisciplinares mediante la construcción de elementos primarios de las infraestructuras.

Aborda tres problemas diferenciados. En primer lugar, presta apoyo al estudio de problemas ambientales complejos, por ejemplo: el cambio climático. Las cuestiones ambientales representan sistemas adaptativos complejos que pueden afectar a muchas disciplinas diferentes y requerir conocimientos de diferentes áreas académicas.

El segundo problema es la falta de uso de formatos compatibles para almacenar datos. Existen retos adicionales en relación con los datos, como puede ser la pérdida de datos (desastre natural, la obsolescencia del formato, los datos huérfanos), las fuentes de datos dispersas, diluvio de datos (el flujo de datos cada vez más heterogéneos), las prácticas deficientes de datos, así como su durabilidad.

El tercer inconveniente es la necesidad de abordar un problema global con una perspectiva global, pues hacer frente a estas cuestiones de forma parcial no es suficiente.

Está diseñado para encarar la necesidad de acceso a datos, seguro y sólido, fundamental para que los esfuerzos de investigación sean productivos y para que puedan elaborarse políticas sobre cuestiones ambientales. Es una plataforma colaborativa e innovadora para la gestión de la ciencia medioambiental que utiliza ciberinfraestructuras y una red de nodos distribuida para proporcionar acceso a datos de observación de la Tierra abiertos, robustos y persistentes de forma segura. Está dentro de la estructura de DataNET.

Existen tres nodos de coordinación (la Universidad de Nuevo México, la Universidad Santa Bárbara de California y la Universidad de Tennessee) y, después, nodos miembros distribuidos geográficamente. La arquitectura técnica está diseñada para permitir a los científicos crear y depositar los datos descritos por cualquiera de las docenas de estándares de metadatos y formatos de archivo que, hoy, se usan con carácter general.

Existe, por consiguiente, una diferencia fundamental en la forma de gestión de las e-infraestructuras en Europa y en Estados Unidos. Mientras que en Europa están sujetas a un régimen jurídico establecido (ERIC), en Estados Unidos se hace mediante subvenciones *ad hoc*.

E.2. NEON

F. Otros modelos continentales

Existen, adicionalmente, otros modelos continentales a los que se hará una breve referencia: el South African Biodiversity Institute (SANBI), el Centro de Referencia en Informação Ambiental (CRIA), de Brasil, The Atlas of Living Australia (ALA) o las e-infraestructuras puestas en marcha por la Academia China de Ciencias (CAS). Aunque distintos del modelo europeo y del estadounidense, estas e-infraestructuras han formado parte de un proyecto financiado en el marco de la Unión Europea.

El South African Biodiversity Institute (SANBI) es uno de los que merece una consideración especial. Este instituto es el responsable de liderar y coordinar la investigación y de supervisar e informar sobre el estado de la biodiversidad en Sudáfrica.

Después de Indonesia y Brasil, Sudáfrica es uno de los países con mayor nivel de diversidad biológica en el mundo. Aunque ocupa menos del 2 % del territorio de la Tierra, representa el 10 % de las plantas de la Tierra, el 7 % de los reptiles, pájaros y mamíferos y el 15 % de las especies marinas costeras conocidas. Dispone de nueve biomas (paisajes únicos de vegetación), de los que tres han sido declarados puntos de acceso de la biodiversidad global.

Persigue construir puentes entre la ciencia y el conocimiento y su implementación. Thabo Mbeki, entonces presidente de Sudáfrica, creó esta organización en 2004 mediante una ley dictada por la National Environmental Management: Biodiversity Act 10 of 2004.

Tiene un mandato legal para gestionar la información de biodiversidad. Una gestión eficiente de la información sustenta la investigación, la toma de decisiones, el asesoramiento de las políticas y la supervisión. Está integrado por una serie de organizaciones y está extendiendo su ámbito de influencia con la intención de ser la referencia de los países de África del Sur.

La Dirección de Información y Planificación de la Biodiversidad es la responsable de la movilización y publicación de datos de biodiversidad para que los usuarios puedan acceder libremente a los mismos. Esta e-infraestructura moviliza información primaria de biodiversidad para ponerla a disposición de los investigadores, de forma gratuita, en Internet con la finalidad de que los que realizan las políticas, los gestores y los investigadores puedan tomar decisiones informadas que contribuyan al desarrollo sostenible de Sudáfrica. Lo hacen, fundamentalmente de cuatro maneras. En primer lugar, promueven el intercambio de datos e información mediante el establecimiento de unos estándares:

- Desarrollando y manteniendo meta bases de datos.
- Estableciendo y manteniendo un portal web.
- Promoviendo las subvenciones relacionadas con la digitalización de datos.
- Dando subvenciones para interfaces de proveedores de datos.
- Facilitando el desarrollo de herramientas y aplicaciones para el uso de los datos.
- Promoviendo la coordinación e interconexión de actividades.

En segundo lugar, contribuyen a la enseñanza, formación y desarrollo de la capacitación con el objeto de promover el acceso nacional a los datos. En tercer lugar, facilitan la concienciación y divulgación de los beneficios y aplicaciones para los usuarios finales. Por último, asesoran en relación con actividades nacionales e internacionales. Funciona como una red de redes de datos, una meta-infraestructura que permite acceder a una multiplicidad de e-infraestructuras, es decir, es una integración.

El Centro de Referência em Informação Ambiental (CRIA), de Brasil, es una organización privada, sin ánimo de lucro. Creado en el año 2000, en el 2002 el Ministerio de Justicia de Brasil la declaró como una organización civil de interés público para la sociedad.

Contribuye directamente a la conservación y utilización racional de la biodiversidad en Brasil. El objetivo fundamental de esta organización es desarrollar la e-infraestructura brasileña de datos de biodiversidad. Además del esfuerzo continuado

por mejorar la infraestructura de datos de biodiversidad, está ahora enfocado en el desarrollo y la implementación de aplicaciones y de servicios de último nivel para dar soporte a la diseminación y el análisis de datos de biodiversidad.

Como consecuencia de ello, representa una gran contribución para la diseminación y publicidad de los datos de biodiversidad y el hecho que se haya convertido en un referente, no solo a nivel nacional, sino también internacional, en lo que respecta al tratamiento de datos de biodiversidad de Brasil. Además, es clave en la promoción del intercambio de datos de biodiversidad.

The Atlas of Living Australia (ALA) es una e-infraestructura pública creada por el Gobierno australiano a través de Estrategia de Infraestructuras de Investigación Colaborativas (NCRIS, por las siglas en inglés de National Collaborative Research Infrastructure Strategy).

Una de las mayores barreras en Australia para la investigación y gestión de la biodiversidad ha sido, históricamente, la fragmentación y la imposibilidad de acceder a algunos datos de biodiversidad. La información podía encontrarse en los herbarios, museos, universidades, organizaciones y departamentos del Gobierno.

Como consecuencia de ello, la obtención de los datos era costosa en tiempo y dinero y, además, con mucha frecuencia resultaba en información incompleta. Por eso, el Gobierno australiano decidió construir una herramienta colaborativa con la finalidad de intercambiar y analizar la información de biodiversidad de Australia.

El objetivo principal de ALA es, por lo tanto, hacer que la información de biodiversidad esté accesible y sea susceptible de utilizarse. Basada en principios de intercambio de datos —recoger los datos una única vez, compartirlos, usarlos todas las veces que sea necesario—, proporciona acceso en línea y libre a más de sesenta y siete millones de registros (especímenes de colecciones de historia natural, campos de observación, etc.).

Este repositorio de información es el más amplio y accesible conjunto de datos de la biodiversidad en Australia. Abre nuevas posibilidades de investigación y la hace accesible a través de Internet. Adicionalmente, es el nodo del GBIF en Australia. Los beneficios de combinar la colaboración internacional con las tecnologías digitales suponen una mejora de los datos entregados y del acceso, con unos costes y riesgos más bajos.

La Academia China de Ciencias (CAS) se creó en Pekín, en 1949. Es el eje, el motor de China, en lo que respecta al estudio y la promoción de la alta tecnología y de las ciencias naturales para el beneficio de China y del resto del mundo.

Dispone de una amplia red de desarrollo e investigación. Integra más de ciento veinticuatro instituciones, con ciento cuatro instituciones, cinco universidades y organizaciones de soporte, doce organizaciones de gestión donde se encuentran las sedes centrales y filiales y otras tres unidades.

Reúne a científicos e ingenieros tanto chinos como de otros lugares del mundo para resolver problemas teóricos y aplicados. Juega una multiplicidad de roles liderando la innovación tecnológica, proporcionando espacios de pensamiento, buscando jóvenes talentos, etc.

Como respuesta a la intención de China de poner la innovación en el centro del desarrollo del país, elabora una estrategia que hace hincapié en la confianza en la gestión democrática, en la apertura y en la promoción de la innovación en la investigación.

G. Coordinación para la realización de análisis globales en la escala espacial y temporal

G.1. CREaTIVE-B

La creación de foros hace posible la elaboración de *roadmaps*; a su vez, la construcción de estos favorece el trabajo conjunto: los programas de planificación y modelización. La finalidad es traer el análisis de cómo puede cristalizarse la aplicación de este nuevo paradigma a las e-infraestructuras de datos de biodiversidad.

El primero de los *roadmaps* se elabora en el marco de un proyecto realizado dentro de los Programas Marco de la Unión Europea (dentro del 7.º Programa Marco), programas de subvenciones, denominado «Coordination of Research e-Infrastructures Activities Toward an International Virtual Environment for Biodiversity» (CREaTIVE-B). Analiza la aplicación práctica a un caso de estudio, sin limitarse a su impacto en la Unión Europea. En efecto, reúne *inputs* de distintos modelos de infraestructuras continentales (LifeWatch, Data ONE, ALA, SANBI, CRIA, CAS).

Recorre, también, a infraestructuras de ámbito más global (GBIF). El centro de este proyecto es la reutilización de los datos. Con esa finalidad trabaja en la identificación del tipo de barreras que obstaculizan no solo el acceso, sino también la reutilización de datos entre las distintas infraestructuras.

El uso de un caso de estudio permite analizar en detalle el tipo de barreras: técnicas, semánticas y jurídicas, y las maneras de superarlas. El tercer *deliverable* analiza las barreras técnicas y semánticas. En el cuarto, se aborda lo que afecta a la interoperabilidad legal.

Todas las e-infraestructuras son públicas, salvo CRIA, que es privada. No existe mandato de aplicación de ningún tipo de estándar de dato ni de metadato, ni restricciones a la hora de usar estándares internacionales.

Los usuarios de las e-infraestructuras tienen la obligación de aceptar los términos de uso de la misma antes de empezar a usarla y la libertad para regularse. Todas las e-infraestructuras están trabajando en programas de formación y uso de estándares comunes. Las condiciones de utilización y reutilización de los datos se fijan en las subvenciones y *grants*, que han dado lugar a la creación de la e-infraestructuras.

En su mayoría, los que subvencionan están abiertos a ajustar sus condicionantes a procesos que promuevan la interoperabilidad, fundamentalmente cuando eso permita llegar a otras comunidades científicas y mejore la posibilidad de abordar el acceso a datos regionales o globales. La mayor parte de las comunidades de usuarios de las e-infraestructuras consideran que las políticas de control de calidad de gestión de estructuras de datos tienen que ser una prioridad.

El análisis de las estructuras legales de las e-infraestructuras lleva a la conclusión de que no existen limitaciones jurídicas ni de ninguna otra índole que impidan elaborar un marco de colaboración a largo plazo entre todas ellas. Sin embargo, compartir datos de biodiversidad plantea cuestiones más complejas que otro tipo de datos científicos. En efecto, pueden estar limitados como puede ser el caso de especies en peligro de extinción; por ello los datos de biodiversidad pueden ser muy sensibles.

G.2. GLOBIS-B

Terminado este proyecto y, tomando como base sus conclusiones, la Unión Europea, en el marco del Programa de Subvenciones Horizonte 2020, da un paso más y auspicia GLOBal Infrastructures for Supporting Biodiversity Research (GLOBIS-B). Desde este proyecto, se quiere contribuir a la creación de las variables esenciales de biodiversidad, con la colaboración de las mismas e-infraestructuras y en un marco de colaboración con el GEO BON.

La existencia de datos permite realizar una mejor monitorización de la biodiversidad. Como ya se puso de relieve al tratar el CDB, la inexistencia de indicadores para medir la pérdida de biodiversidad dificulta frenar su pérdida. El hecho de que existan datos no es suficiente, sino que es necesario elaborar y construir indicadores como los que existen en el marco del IPCC. El GEO BON empieza a construir el concepto de variables esenciales de biodiversidad y el GLOBIS-B da servicio al GEO BON, por lo que contribuye a la creación de los indicadores.

La finalidad de las EBV es ayudar a las comunidades de observación a armonizar la monitorización identificando las variables que es necesario medir. Existen unos

criterios de elaboración de las EBV: idealmente, tiene que ser capaz de capturar escalas y dimensiones de biodiversidad críticas; tienen que ser biológicas y una variable de estado (con carácter general); es necesario que, en la medida de lo posible, sean agnósticas al ecosistema (es decir, que podrán usarse en cualquier ecosistema, marino, de agua dulce, etc.) y, por último, deben ser técnicamente factibles, viables económicamente y sostenibles en el tiempo. Además, deben poderse medir y modelar globalmente. Una característica fundamental de las EBV es que se encuentran entre la observación primaria y los indicadores.

Su objetivo es identificar los datos primarios requeridos, las herramientas de análisis, las metodologías y los cuellos de botella jurídicos y técnicos para desarrollar una agenda de desarrollo de la investigación y de las e-infraestructuras para programar las EBV.

Por lo tanto, las EBV se definen como una medida necesaria para estudiar, reportar y gestionar el cambio en biodiversidad y, potencialmente, en esta definición encajan cientos de variables. Por esta razón, se está trabajando en un análisis con la finalidad de identificar las que cumplen los criterios de escalabilidad, sensibilidad temporal, viabilidad y relevancia.

Es notoria la falta de coordinación entre sistemas que hace que los datos y conjuntos de datos procedentes de distintas fuentes sean irreconciliables. Las EBV quieren dar respuesta y solución a esta descoordinación mejorando el acceso y el intercambio de los datos.

La implementación de las EBV tiene que abordar dos retos fundamentales: los de carácter técnico —las cuestiones de interoperabilidad— y los de carácter científico —relacionados con el propio concepto de EBV, es decir, cómo se definen exactamente las EBV, qué datos de biodiversidad se necesitan, la disponibilidad de los mismos y de dónde se obtienen, de qué manera se puede acceder a los datos relevantes, cómo se pueden calcular las EBV, cuáles son los ámbitos espaciales, temporales y actuales relevantes y cómo de sensibles son las EBV a las variaciones que se produzcan en los datos en los que se basan—.

Para analizar la magnitud y la dirección del cambio de las dimensiones esenciales de la biodiversidad, el enfoque que propone el GEO BON es utilizar los distintos niveles de organización biológica: genética, taxonomías, funciones y estructura. Con arreglo a estos distintos niveles, se han construido las clases de EBV.

Por el momento, el GEO BON ha propuesto veintidós EBV agrupadas en seis clases: composición genética, poblaciones de especies, características de especies, composición de comunidades, funcionamiento de ecosistemas y estructura de ecosistemas. Aunque se han empezado a dar los primeros pasos, y parece que se va a buen término, todavía está en proceso de elaboración.

En el mundo, tal como estaba concebido antes, se vivía y se investigaba a un nivel más local, que parecía más sujeto a las jurisdicciones o normas nacionales. Sin embargo, de un tiempo a esta parte, parece que el mundo se ha hecho más grande y las normas nacionales parece que ya no son suficientes. Por eso, aparecen algunas cuestiones relacionadas con la posibilidad de compartir, de reutilizar, etc.

El análisis de los *roadmaps* lleva de lleno a la cuestión de la interoperabilidad legal, esta vez acotada en el marco de los dos *roadmaps*. En el marco de CREaTIVE-B, la atención está centrada en ver cómo se puede construir una red de bases de datos de distinto ámbito nacional (NEON, DataONE, ALA, CRIA, CAS, etc.), global (GBIF) y supranacional (LifeWatch) y hacer que trabajen conjuntamente, que puedan colaborar. El arranque de la interoperabilidad se hizo en este proyecto desde el análisis de cada uno de los contextos de las bases de datos para poder detectar posibles barreras a la interacción.

De partida, las cuestiones que se plantearon fueron si podía haber normas de base (las que sirvieron de fundamento a la creación de las distintas bases de datos) que obstaculizaran la apertura de los mismos. De igual manera, si existía obligación o restricciones a la hora de compartir datos, etc. Después de un análisis detallado de cada una de las regulaciones a que estaba sujeta cada una de las bases de datos, resultó que todas las e-infraestructuras trabajan sin restricciones en relación con la reutilización de datos. Aunque de base establecen políticas de acceso abierto de datos, todas las e-infraestructuras tienen una política pragmática en lo que respecta a la reclamación sobre propiedad de datos, separándose del síndrome *it's my data*. Ciertamente, el *roadmap* reconoce que pueden existir restricciones, por ejemplo, como consecuencia de los derechos de propiedad intelectual o *copyright*, o también restricciones derivadas del uso de *software* que no sea de código abierto. Otra de las posibles restricciones que planea sobre esto es la aprobación de la directiva europea sobre minería de texto y datos.

A efectos de interoperabilidad, se dan una serie de recomendaciones. Una sería abrir lo máximo posible los datos a la reutilización de los mismos; otra, el uso de las licencias CC (fundamentalmente en sus formas CC4 y CC0, hoy, sin perjuicio de que puedan aparecer otras licencias fáciles de usar que sirvan para abrir los datos y facilitar su reutilización); otra, fomentar que el procesado de datos pueda hacerse de forma automática; otra, por supuesto, trabajar en datos de calidad. Existen otras, aunque las mencionadas son las más relevantes.

Por último, se abordan las cuestiones de interoperabilidad en el proyecto de GLOBIS-B. En este caso, el enfoque del proyecto, superada la primera fase de interoperabilidad básica entre bases de datos, está en las EBV; por lo tanto, la interoperabilidad necesita dar un paso más. En efecto, la problemática que plantean las EBV tiene una complejidad mayor, puesto que el tipo de interoperabilidad tiene

que ser casi total. Las variables sobre las que se trabaja son muy numerosas y cada vez lo serán más. Si el concepto consigue construirse, los datos que alimentarán las mediciones serán casi de todas las fuentes conocidas. Cuanto más completo sea, más completa será la medición.

En este contexto, vuelven a plantearse todas las cuestiones analizadas a lo largo de este trabajo y que exigen dar un paso más. ¿Es una forma de poner otra vez todo en tela de juicio como excusa para dar un salto cualitativo en el diseño de la interoperabilidad legal global?

ANEXO 1: CONCLUSIONES DEL PROYECTO GLOBIS-B

GLOBIS-B (654003)



D4.3: Final position paper for policy makers

Project acronym: *GLOBIS-B*

Project full title: "GLOBal Infrastructures for Supporting Biodiversity research"

Grant agreement no.: 654003

GLOBIS-B (654003)

Due-Date:	31 May 2018
Actual Delivery:	31 May 2018
Lead Partner:	UAH
Dissemination Level:	PU
Status:	Public version
Version:	0.3

GLOBIS-B (654003)

DOCUMENT INFO

Date and version no.	Author	Comments/Changes
23 May 2018, v0.1	Wouter Los, UvA	Structuring and outlining of content. Draft for project approval.
30 May 2018, v0.2	Wouter Los, UvA	Edits.
30 May 2018, v0.3	Enrique Alonso, UAH	Final check and corrections.

GLOBIS-B (654003)

TABLE OF CONTENTS

1. Executive summary	6
2. Contributors	6
3. Background	7
3.1. Operationalizing the EBV framework	8
3.2. Requirements for making EBVs available for policy purposes	8
3.3. Workshop agenda and methodology	9
4. Workshop outcomes	10
4.1. EBVs as a fundamental resource for the evaluation of biodiversity trends	10
4.2. Agreed recommendations	11
4.3. Policy recommendations	12
4.4. Recommendations for Research Infrastructures	14
5. Conclusions	16
6. References	16
Annex 1 Participants of the policy workshop	17
Annex 2 Information package as sent to workshop participants	18

GLOBIS-B (654003)

LIST OF FIGURES

Page 10. Figure 1: EBVs are calculated from raw data and provide building blocks for indicator development.

Page 21. Figure 2: Essential Biodiversity Variables as positioned between raw data and indicators.

GLOBIS-B (654003)

1 Executive summary

Exploiting primary biodiversity and other data to produce and manage Essential Biodiversity Variables (EBV) data products depends on cooperation, practicality and interoperability among multiple stakeholders, including those collecting and mobilising data with EBV potential (EBV-usable data), those making data 'EBV-ready' and those producing, publishing and preserving EBV data products. Since EBVs feed the construction of different indicators for policy use, also policy makers are important stakeholders to provide guidance on priorities. A policy workshop organised by GEO BON and the GLOBIS-B project brought in May 2018 together the representatives of various (regional and international) policy bodies with senior officials of biodiversity research infrastructures from different continents. The workshop addressed a number of findings:

- While building EBV data products requires a considerable level of data integration from a large number of data providers, it is possible to run comparable, repeatable and quality-based EBV workflows. Policymakers and funders could help move EBVs beyond an academic exercise by expressing their views about the further operationalization of EVBs for their use in biodiversity change assessments and about the role of EBVs in current and future policy developments.
- The research infrastructures contributing to GLOBIS-B have expressed their interest to cooperate on EBV production at the global scale. Policymakers and funders could help promoting coordination around EBV production, including the development of shared and repeatable workflows, and encourage infrastructures to make EBV data products open as well as findable, accessible, interoperable, and reusable (FAIR).
- EBV production is often constrained by legal restrictions on the use, modification and sharing of software and data. Policymakers and funders could help clarify and incentivize requirements for legal interoperability, emphasizing the value of automated and unrestricted data retrieval, and the use of standardized, machine-readable copyright waivers such as CC0, or licenses such as CC-BY.

The workshop discussions resulted in a clear set of recommendations for both the policy sector and the research infrastructures. GEO BON will bring the policy recommendations to the attention of policy bodies, and with clear suggestions for their international meetings. The cooperating research infrastructures in the framework of the GLOBIS-B project will continue on the recommendations relevant for their operations.

2. Contributors

The authors of the present document are Enrique Alonso Garcia (WP4 leader and UAH) and Wouter Los (UvA), while many other members of the GLOBIS-B project team and the participants of the final GLOBIS-B policy workshop provided contributions shaping the outcome of the work. The list of workshop participants is in Annex 1 of this document. We acknowledge the assistance and contributions of the following individuals who contributed significantly throughout the project to the continuing discussions about legal and policy issues related to the development of EBV data products: Donat Agosti, Anne Bowser, Willi Egloff, Laetitia Navarro, and Paul Uhler.

GLOBIS-B (654003)

3 Background

Although reducing the rate biodiversity loss and averting dangerous biodiversity change are accepted international goals, there is no global, harmonized observation system for delivering regular, timely data on biodiversity change¹. A key mechanism for studying and reporting on biodiversity change is the concept of Essential Biodiversity Variables (EBVs). EBVs are a minimum set of critical variables representing different dimensions of biodiversity from genetic composition, species populations, species traits, community composition, ecosystem function, to ecosystem structure. Processed EBV datasets for each dimension (class) provide the basis for developing high-level indicators of biodiversity. GEO BON² is improving the acquisition, coordination and delivery of biodiversity observations and related services to users, including decision makers and the scientific community. A global biodiversity observation network should contribute with data to effective management policies for the world's biodiversity and ecosystem services.

The approach of Essential Biodiversity Variables (EBVs) as developed by GEO BON has been analyzed by the EU-funded GLOBIS-B project³ in order to propose standardized computational workflows for delivering EBVs. The project has studied how international cooperation of research infrastructures can support standardized EBV development through data sharing and by offering computational workflow tools. Bringing this into effect depends on the willingness of the infrastructures and their governing bodies to contribute with each other's strengths to cooperative efforts in support of EBV production. But it also depends on the ability of national and international policy bodies to provide guidance on priorities, and on removing legal and financial barriers for cooperation.

The GLOBIS-B⁴ project "Global Infrastructures for Supporting Biodiversity research" is a Horizon 2020 funded initiative within the coordination and support funding scheme for international research infrastructure cooperation of the European Commission. The project brings together scientists, IT and legal experts, and operators of research infrastructures to address the research needs and infrastructure services required to support EBVs. This cooperation has demonstrated, for a few EBV classes, how the harmonization of data collection and technical data management can contribute to standardized and repeatable EBV data products with common computational workflows. EBVs as produced with data from various sources (e.g. in situ monitoring, remote sensing) are crucial to assess changes in biodiversity through time. They can be used to measure the achievement of policy goals such as the Aichi Biodiversity Targets set by the Convention on Biological Diversity (CBD), or the UN Sustainable Development Goals (SDGs), or the national targets defined in National Biodiversity Strategies and Action Plans (NBSAP) or the like. Finally, they can also serve to define biodiversity management policies from local to global scales.

¹ Pereira, H. M., Ferrier, S., Walters, M., Geller, G., Jongman R. H. G., Scholes, R. J., Bruford, M., Brummitt, N., Butchart, S. H. M., Cardoso, A., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D., McGeoch, M. A., Obura, D., Onoda, Y., Pettoirelli, N., Reyers, B., Sayre, R., Scharlemann, J. P.W., Stuart, S. N., Turak, E., Walpole, M. & Wegmann, M. (2013). Essential biodiversity variables for global earth observation. *Science*, 339: 277-278. DOI: 10.1126/science.1229931

² <http://geobon.org>

³ GLOBIS-B: GLOBal Infrastructures for Supporting Biodiversity research. EU-H2020 grant agreement No 654003.

⁴ <http://www.globis-b.eu>

GLOBIS-B (654003)

3.1. Operationalizing the EBV framework

Building EBV data products with heterogeneous types of data in the form of time series requires identification and clear definition of the key dimensions (space, time and taxonomy), attributes and uncertainties of EBV-useable data. Such data must comply with minimum requirements to be defined and be based on reliable data that have been collected and processed following scientifically recognized standards. The data sources and the construction of EBV data products must be traceable and the processing method must be reproducible.

To document biodiversity changes comprehensively, EBVs depend both on existing data, and on new data generated in the future through targeted monitoring. EBVs should be flexible and relevant, and should be possible to calculate:

- For any geographic area,
- At any spatial and temporal scale,
- For any species, assemblage, ecosystem, or biome of interest,
- With data held by any (or across) data repositories, and
- By any expert willing to deploy the conceptual and operational framework of EBVs.

As EBVs will usually be built from multiple sources and research / monitoring activities, combining them requires sufficiently informative metadata about the EBV usable data to allow harmonization and standardization. Building EBV data products further requires dealing with different scales, correcting for imperfections, interpolation and extrapolation; as well as quantifying uncertainties. EBV data products should follow a standardized structure to be defined and the approaches of EBV production should be compatible. The conclusions of the GLOBIS-B project for constructing EBV data products for species distribution and abundance (EBV class species populations) has been published in Biological Reviews 2018 (<https://doi.org/10.1111/brv.12359>).

3.2. Requirements for making EBVs available for policy purposes

If EBV development is to transcend scientific curiosity to become available for policy purposes, it will require clarity and support with respect to:

- A. Policy priorities on required EBV products (at the levels of prioritized species, assemblages, ecosystems, biomes, areas, scales);
- B. Coordinated monitoring schemes for primary data collection/production on required scales in support of constructing EBV-usable datasets, such as foreseen with Biodiversity Observation Networks (BONs) around the world;
- C. Computational workflows to process primary data into various EBV data products;
- D. Cooperation of research infrastructures for curating processed (EBV) data and publishing them in required formats;
- E. Overcoming legal constraints in relation to accessing data and achieving workflow interoperability.

All workshop participants received prior to the workshop a background document providing considerations with respect to the above issues A to E, together with information about information GEO BON and the GLOBIS-B project (Annex 2).

GLOBIS-B (654003)

3.3. Workshop agenda and methodology

The workshop agenda below covered Tuesday afternoon 8 May 2018 and almost the full day of 9 May. After introductory presentations, the workshop participants were divided in four groups to discuss implications of the EBV requirements as mentioned in paragraph 3.2. Afterwards, the group reports allowed to structure suggested recommendations under a few headlines. Next, these structured recommendations were further discussed in one group on policy issues, and another one on issues relevant for research infrastructures. A plenary review of these discussion outcomes resulted in a concise set of recommendations, supported by all participants.

Agenda

8 May 2018

1. 13:30 Welcome & Introduction to the objectives of the Workshop
2. 14:00 Keynote lectures (*Chair: Daniel Kissling*)
 - EBVs as key input for policy objectives (*Laetitia Navarro*).
 - Results from the GLOBIS-B project relevant for the workshop (*Daniel Kissling*).
 - Test case (ALA & GBIF) on invasive species (*Alex Hardisty*).
 - EuroGEOSS – Coordinate, combine and cooperate (*Marjan van Meerloo, European Commission*)
- 15:00 Coffee / Tea
3. 15:30 – 17:30 Table groups of 6 persons each will discuss the proposed recommendations.
 - A. Policy priorities on required EBV products (species, assemblage, ecosystem, biome, areas, scales).
 - B. Coordinated monitoring schemes for primary data collection or production on of constructing EBV-usable datasets, such as foreseen with Biodiversity Observation Networks (BONs) around the world.
 - C. Computational workflows to process primary data into various EBV data products.
 - D. Cooperation of research infrastructures for curating processed (EBV) data products and publishing them in required formats.
 - E. Overcoming legal constraints in relation to accessing data and achieving workflow interoperability.

Discuss and agree on each of the recommendations, note any special remarks, and highlighting (if necessary) any significant points of concern.
4. 17:30 – 18:00 Presentations by table group rapporteurs (*Chair: Laetitia Navarro*)

9 May 2018

5. 09:00 Discussion outcomes and proposed conclusions (*Wouter Los*).
6. 09:30 Initial discussion on proposed conclusions from the workshop with a list of recommendations (*Chair: Laetitia Navarro*).
- 10:30 Coffee / Tea
7. 11:00-13:15 (Continued) Discussion proposed conclusions from the workshop with a list of recommendations (*Chair: Laetitia Navarro*).
- 13:15 Lunch and end of workshop
8. 14:00 – 15:30 Publicity and next steps.

Round table discussion about next steps.

 - What possible actions for agreed recommendations?
 - Who will act on each of these?
 - Publicity on workshop outcomes.

D4.3 Final position paper for policy makers

Page 9 of 24

GLOBIS-B (654003)

GLOBIS-B (654003)

4. Workshop outcomes

4.1. EBVs as a fundamental resource for the evaluation of biodiversity trends

Policy questions and needs are varying over countries, regions, continents and globally, and these change over time. Depending on existing and/or emerging policy questions, indicators need to be defined that are informative for answering the expressed questions and needs. As such, preferred indicators will likely also change over time. In order to avoid that new data have to be generated for each new indicator, it is attractive to build upon a body of EBV data sets. Depending on any preferred indicator, it should be possible to construct the indicator from a data selection in a full body of EBV data sets. In this view, a body of EBV data sets should be regarded as a stable intermediate data layer in between *raw data* and varying *indicators*. Examples include EBVs to support infectious disease vectors, invasive species, and effects of biodiversity management. This is needed to guide national-level initiatives among producers and consumers who can then contribute to global-level dialogue. Moreover, sufficiently large EBS data sets would allow for predictive ecology: potentially forecasting trends and assessing the impact of management interventions.

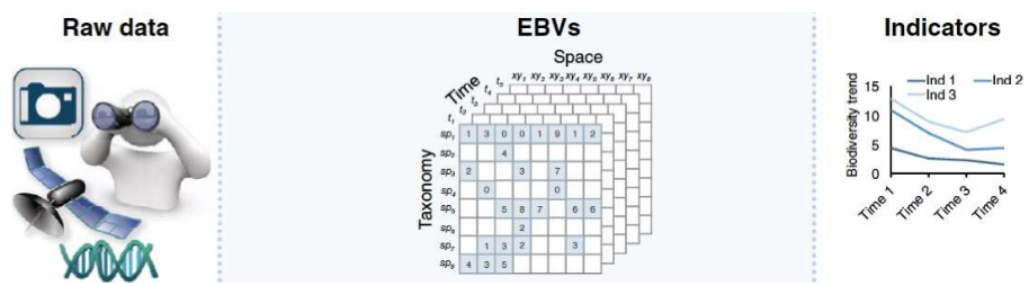


Figure 1: EBVs are calculated from raw data and provide building blocks for indicator development.

To facilitate varying indicators, it is the challenge to agree on building a body of sufficiently large EBV data sets, that are if possible are based upon cooperation, practicality and interoperability among multiple stakeholders, including those collecting and mobilizing data with EBV potential (EBV-usable data), those making data 'EBV-ready' and those producing, publishing and preserving EBV data products. Ten principles offer best current practice guidance to data and infrastructure organizations to enhance their ability to contribute towards production of global EBV data products, whilst retaining autonomy and flexibility to achieve what is needed in ways appropriate to the organizations' own business. The principles as developed in agreement with the cooperating research infrastructures are summarized below.

1. Data Management Plan

Projects developing EBV data products should have comprehensive data management plans.

2. Data Structure

EBV data products should adhere to agreed-upon minimum dimensions for each product (i.e., time, space, name/taxonomy (where applicable), etc.). All EBVs should be accommodated in a common framework that conforms (as far as is practically possible) with content and schema standards for representation formats and exchange protocols.

3. Metadata

EBV data products and the EBV-ready datasets from which they are generated should have associated human- and machine-readable metadata, compliant with accepted community standards and sufficient for purposes of data discovery, access, fitness-for purpose evaluation, citation, interpretation and use.

GLOBIS-B (654003)

4. Services

EBV data products, EBV-ready datasets, digital objects and other related services should expose their capabilities and be accessible through common, standardized Application Programming Interfaces (APIs).

5. Data Quality (Fitness-for-use)

Each EBV data product and EBV-ready dataset should include data quality documentation, sufficient to identify fitness-for-use of the data for specific purposes.

6. Workflows

It should be possible to execute published, standard workflows for preparing, publishing and preserving EBV data products and the EBV-ready datasets from which they are produced. Ideally, such workflows should be defined and represented in a non-proprietary manner.

7. Provenance

It should be possible to trace the EBV production process from the product back to the primary data and to reproduce the process. Provenance information must be readable both by humans and by machines.

8. Ontologies / Vocabularies

EBV data products and EBV-ready datasets should be described by standard, openly accessible and machine-readable key vocabulary terms and conceptual relations (ontologies) presented in a simple way to enable wide usage.

9. Data Preservation

EBV data products and associated underlying data should be preserved with an associated persistent identifier in a community supported and trusted repository.

10. Accessibility

EBV data products and EBV-ready datasets must be sensitive, timely and FAIR (Findable, Accessible, Interoperable and Reusable).

4.2. Agreed recommendations

The recommendations below as resulting from the workshop discussions are focussing on four main areas:

1. Secure trust in a stable body of EBV data sets
2. Put in place the conditions for effective EBVs fit-for-use
3. Biodiversity research infrastructures around the world can make this work
4. Advertise the recommendations and request policy support

The recommendations are split in ones addressing policy bodies, and others for research infrastructures.

GLOBIS-B (654003)

4.3 Policy recommendations

Policy recommendations are for audiences at the national, regional and international level. Although national policy levels are mainly acting in legislation and management, the bodies based on international agreements and conventions also are suggested to consider the recommendations below.

Secure trust in a stable body of EBV data products

- Policy bodies are recommended to bring into promote that EBV data sets will be based on traceable and quality controlled primary data. Feedback mechanism(s) are needed from uses to providers.
- Building trust and sustainability across entire value chain by ensuring appropriate funding and appropriate governance, also securing appropriate attribution for contributors.
- Likewise, EBV data products must be open and FAIR, and funding bodies are recommended to request this in their funding conditions.
- Trust also requires a professional approach with a competent research community, together with appropriate mechanisms to review EBV development. Here funding bodies can promote such mechanisms.
- The long-term sustainability of stable EBV data products should be implemented by the research infrastructures concerned, and their funding authorities are recommended to explicitly require this.

Put in place the conditions for effective EBVs fit-for-use

- Policy stakeholders are invited to consider improved indicators through the availability of EBV products. They are also suggested to engage with the communities generating the required primary data and the ones constructing EBVs to promote and ensure that the process from primary data to indicators is streamlined.
- Generating primary data relevant for different EBV classes is not yet comprehensively organized with standardized protocols to address data gaps. Funding bodies are recommended to provide funding as incentive to accelerate this.
- Recommend that legislation is effective for opening up and sharing data, and for cross-border data use in the countries and economies where it is still not yet fully implemented.

Advertise the recommendations and request policy support

- GEO BON is invited to take up the heritage of the GLOBIS-B project and to advertise the recommendations in this paragraph.
- A strategy is required for including EBVs in the agenda of the post-2020 biodiversity strategy. The fifteenth meeting of the Conference of the Parties in 2020 is expected to update the Convention's strategic plan. In the context of the 2050 Vision of the current Strategic Plan for Biodiversity 2011-2020 as well the 2030 Agenda for Sustainable Development and other relevant international processes, and in the light of an assessment of progress in achieving the goals and Aichi Biodiversity Targets of the current plan as well as of future scenarios of change. Since the deadline to submit documents or relevant input to the process of preparations for Post-2020 Biodiversity Framework, within the phase of proposals for a comprehensive and participatory process is closed since 16 February 2018, (See <https://www.cbd.int/post2020/>) and, following the peer review process the document will be revised and made available for consideration by the Subsidiary Body on

GLOBIS-B (654003)

Implementation SBI) during its second meeting that will take place in Montreal, Canada, 9 - 13 July 2018 and the documents to be analyzed by the SBI seems also closed (see <https://www.cbd.int/sbstta22-sbi2/review.shtml#tab=1>), the first next step is to analyse how the conclusions and potential additional relevant documents related to GLOBIS-B can be included either info documents or during its celebration through delegates of the countries with which the Research infrastructures have connections as well as the European Commission, or through observers attending its meeting (i.e. GEO BON, IPES or GBIF, if they attend). In any case the process toward the final decision will not be closed in this meeting so the participants assume the commitment to closely follow the process until its final approval.

GLOBIS-B (654003)

4.4 Recommendations for research infrastructures

It should be noted that biodiversity research infrastructures have different roles, such as data providers, data aggregators, analytical / computational infrastructure, repositories of EBV data products. In summary the research infrastructures have to consider their position with respect to the three following generic roles:

1. collecting and mobilising data with EBV potential (EBV-usable data),
2. making data 'EBV-ready' and,
3. producing, publishing and preserving EBV data products.

As for the first role, the workshop expressed that data providers deserve much credit. They have already invested much efforts in securing the availability and accessibility of raw data. For the latter third role is nothing yet in place to perform role at present. A general recommendation for the cooperating research infrastructures is to create something new or to encourage existing research infrastructures to take up this role. Users and other stakeholders must be involved in emergence of that and must incorporate feedback mechanism to first two roles.

Secure trust in a stable body of EBV data products

- Policy bodies are recommended to bring into promote that EBV data sets will be based on traceable and quality controlled primary data. Feedback mechanism(s) are needed from uses to providers to build trust and sustainability across entire value chain. It is recommended that research infrastructures offer such mechanism in their operations.
- EBV data products, workflows and applied tools must be open and FAIR. The research infrastructures are recommended to put in place a number of streamlining activities as proposed in the so-called Bari Manifesto (see at the end of Annex 2). These include harmonization around metadata standards, ontologies, and on the level of data records and data quality.
- Engage with relevant standards bodies (such as TDWG, OGC, W3C, RDA, etc.)
- Turn synthesis centres worldwide on to producing the EBV data products they might need to answer the questions they tackle. Engage institutions capable to for example, the 'synthesis centres' to tackle some specific challenges.
- All steps in workflows from primary data to enriched EBV data sets should be recorded, allowing the documentation of everyone's contributions in workflow steps. This assists in tracing successive steps, and so becomes auditable.
- Trust also requires a professional approach with a competent research community, together with appropriate mechanisms to review EBV development.

Put in place the conditions for effective EBVs fit-for-use

- Engage the stakeholders and communities along the EBV value chain in the exercise to construct/compute EBVs data products. These are the experts, for assessing existing data sources, for protocols to generate new data, and for reaching standards.
- Consider a centralized or distributed service with a tool and workflow catalogue (BON-in-a-Box).
- This also should include services to assist in understanding and deploying such services. Generic workflow components of an EBV need to be easily explainable.

GLOBIS-B (654003)

- Give special attention to the training of data scientists capable of operating in this field, from data custodians up to data analytics experts and data executive officers.
- Attention that legislation is effective for opening up and sharing data, and for cross-border data use.

Biodiversity research infrastructures around the world can make this work

- Biodiversity research infrastructures around the world started working together to serve a variety of user audiences. It is potentially possible to construct EBV data products at an industrial scale, and thus build a body of EBV data sets.
- The GLOBIS-B recommendations are addressing a further development from the initial EBV concept, and are suggesting building a body of EBV data sets at an industrial scale allowing for underpinning diverse indicators in support to changing policy and management needs. The research infrastructures are recommended to agree on a common policy for building such EBV data sets at the national and regional levels. This includes an architecture for storing and finding the EBV data sets with options to bring these into use for constructing emerging indicators relevant for policy needs.
- Of course this will start with a few tests, and expanding EBV data sets needs careful joint planning with the coordinating GEO BON.
- A number of data and research infrastructures with the lead of GBIF are discussing the implementation of a shared pipeline for supporting/delivering EBV-ready datasets for the species distribution EBV.
- The LifeWatch research infrastructure discusses developing (with the cooperating Research Infrastructures and the wider scientific community) a shared pipeline with workflows on a dedicated hardware platform to analyze and model the EBV-data.
- These latter processes should be encouraged and the involved research infrastructures are encouraged to promote further collaboration with the involved other research infrastructures and their user communities.

Advertise the recommendations and request policy support

- All recommendations above should be reported to CBD, IPBES and similar conventions in order to promote adoption of this approach.
- Communication needs to be tailored toward audiences. Real life examples may underpin and explain the benefits of the promoted approach. Examples are invasive species, pollinators, algal blooms, and infectious diseases.
- It is important that stakeholders, at national, regional and global scales will recognize the importance of promoting and investing in their areas of capabilities and interests.
- GEO BON is invited to take up the heritage of the GLOBIS-B project and to advertise the recommendations in this paragraph.

GLOBIS-B (654003)

5. Conclusions

The GLOBIS-B resulted in considerable progress in understanding the EBV concept, on how research infrastructures can work together to implement computational workflows for EBV construction, and how both policy authorities and scientific communities may benefit from a resource of EBV data products. A coordinated test on biodiversity change related to invasive species showed that the GLOBIS-B results are feasible in practice. The GLOBIS-B project coordination hereby is expressing gratitude for these efforts by GBIF, the Atlas of Australia and the BioVEL initiative. It is crucial to keep together the scientific, infrastructure, technical, and legal communities that cooperated in GLOBIS-B. More concrete efforts are required to build a body of EBV data products, so that better and scientific based indicators - such as for detecting pollinator change - can be based on these data products.

GEO BON will continue on the momentum created by GLOBIS-B, with publicity and by promoting awareness in the relevant policy and other communities. But also by involving the experts from the GLOBIS-B workshops in its GEO BON working groups. A concrete GEO BON action is to advise (national and international) policy bodies about next steps for EBV implementation, following their earlier support for the EBV approach. The involved research infrastructures will build on this with further tests, and to get a better picture of each one role in the chain from generating raw data, data processing, storing EBV data products, and the provision of services to users including policy. By building on existing efforts, successful examples and explanatory documents/publications will assist the GEO BON community to bring the EBV concept on a next operational level in the next few years.

It is reassuring that GBIF is already discussing the implementation of a shared pipeline for supporting/delivering EBV-ready datasets for the species distribution EBV, and that the LifeWatch research infrastructure is discussing the development of a similar for computational workflows to analyze and model the EBV-data.

6. References

- [Kissling 2015] Kissling, W.D., Hardisty, A., García, E.A., Santamaria, M., De Leo, F., Pesole, G., Freyhof, J., Manset, D., Wissel, S., Konijn, J. & Los, W. (2015) Towards global interoperability for supporting biodiversity research on essential biodiversity variables (EBVs). *Biodiversity*. 2015 Jul 3;16(2-3):99-107. doi: [10.1080/14888386.2015.1068709](https://doi.org/10.1080/14888386.2015.1068709).
- [Kissling 2018] Kissling W D, Ahumada J A, Bowser A, Fernandez M, Fernández N, García E A, Guralnick R P, Isaac N J B, Kelling S, Los W, Mcrae L, Mihoub J B, Obst M, Santamaria M, Skidmore A K, Williams K J, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F, Egloff W, Elith J, Hobern D, Martin D, Pereira H M, Pesole G, Peterseil J, Saarenmaa H, Schigel D, Schmeller D S, Segata N, Turak E, Uhler P F, Wee B and Hardisty A R 2018 Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale *Biol. Rev.* 93 600–25 Online: <http://doi.wiley.com/10.1111/brev.12359>
- [Pereira 2013] Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013) Essential Biodiversity Variables. *Science*, 339, 277-278.

GLOBIS-B (654003)

Annex 1: Participants of the policy workshop (8 - 9 May 2018)

Person last name	First name	Organisation
Agosti	Donat	PLAZI
Alkemade	Rob	IPBES TSU Scenarios and Models
Alonso	Enrique	State Council of Spain
Bubb	Philip	Biodiversity Indicators Partnership
Canhos	Vanderlei	Brazilian Reference Centre on Environmental Information
Egloff	Willi	PLAZI
Hardisty	Alex	Cardiff University
Hirsch	Tim	Global Biodiversity Information Facility
Holewa	Hamish	Atlas of Living Australia
Huis, van	Edwin	Group of global large Natural History Museums
Ji	Liqiang	Chinese Academy of Sciences
Johansson	Anna-Maria	European Commission - Unit Research Infrastructures
Kissling	Daniel	University of Amsterdam
Konijn	Jacco	University of Amsterdam
Koureas	Dimitris	Distributed Systems of Scientific Collections
Krug	Cornelia	Future Earth
Le Bras	Yvan	FR-BON / ECOSCOPE
Leo, de	Francesca	CNR Italy
Londoño	Maria Cecilia	Humboldt Institute Colombia & CO-BON
Los	Wouter	University of Amsterdam
Manset	David	CEO GNUBILA/MAAT
Meerlo, van	Marjan	European Commission - Unit Earth Observation
Michener	Bill	DataONE data observation network
Navarro	Laetitia	GEO BON
Tienderen	Peter H.	LifeWatch research infrastructure

GLOBIS-B (654003)

Annex 2: Information package as sent to workshop participants

The text below presents considerations related to EBV requirements A-E (see page 8). The considerations are directed to political, policy, and funding bodies and to research infrastructures. This is indicated in each sub-section below.

B. Policy priorities on required EBV products (species, assemblage, ecosystem, biome, areas, scales)

The GLOBIS-B project investigated the feasibility of producing EBV data products on biodiversity change, and this provided valuable proof of concept for the EBV framework. To advance the EBV framework from this proof of concept to broader development and uptake, there is a need for top-down policy direction to help researchers and supporting infrastructures understand where to focus subsequent efforts within the variety of possible EBV implementations. Guidance by policy bodies could include information on:

1. The prioritized high level challenges, such as the Aichi Biodiversity Targets or new ones that could be supported by related indicators and relevant EBVs.
2. Any other policy targets that might trigger to take up new dimensions of biodiversity not yet covered in existing indicators.
3. Promoted provision of resources on (sub)national and regional scales for meeting the policy priorities with targeted and large-scale EBV efforts.

B. Coordinated monitoring schemes for primary data collection or production on required scales

The production of EBV data products depends on suitable and accessible data. Historically, before the EBV concept was in place, biodiversity observation and monitoring practices were directed at different objectives, either for research purposes or for other reporting obligations. Parts of such legacy biodiversity data can be used for the construction of EBV data products. But new data collection should follow standardized and structured protocols to serve for EBVs and derived indicators, and international policy guidance can serve national efforts in this regard.

1. the CBD may want to consider how its resolutions on global and national obligations on reporting on biodiversity trends may benefit from the new opportunities for building policies on standardized and scientifically recognized EBV data products⁵;
2. likewise, to establish links with other related conventions;
3. national authorities charged with setting monitoring priorities may want to promote the inclusion of EBV relevant data in adequately resourced monitoring schemes, allowing for harmonized EBV data processing, for which Biodiversity Observation Networks (BONs) are appropriate vehicles;
4. involved stakeholders may exchange information and agree on best practice approaches in this regard. The BON in a box online catalogue could be considered as an adequate platform for capacity building and knowledge exchange (<https://boninabox.geobon.org/>).

C. Computational workflows to process primary data into various EBV data products

It is possible to run comparable, repeatable and quality-based EBV workflows with support from cooperating infrastructures. Ideally, widely accepted and adopted sets of generic workflows could lead to substantial EBV production in support of a range of indicators on biodiversity change. But building EBV data products requires a considerable level of data integration from a large number of dispersed data providers,

⁵ This relates to Aichi Target 19: "By 2020, knowledge, the science base and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied."

GLOBIS-B (654003)

as well as complex preparation and processing steps. This is possible to implement, but research infrastructures and their funders can only provide cost-effective technical actions when policy expectations will provide them with guidance upon which they can base the development of appropriate computational architecture. Considerations are the following:

1. Policy bodies are suggested to consider the relevance of the production of EBV data products. This could be (a) for scientific interests, (b) for national interests, or (c) for wider regional, continental or global comparability and analysis.
2. Similarly, policy bodies may consider whether only ad hoc EBV development is sufficient, or alternatively a factory-scale EBV production should be preferred to support the policy needs. The latter would imply a generic level to re-use approaches globally, rather than a restriction to domestic solutions serving local interests.
3. Stakeholders are suggested to discuss and make recommendations on the responsibilities for each of the steps in the range from priority setting, data collection up to EBV production and maintenance. What should be organized bottom-up and what top-down; what are the implications for governance and finance?
4. What would be the role of GEO BON and its (national regional, thematic) BONs, or any other coordinating body? How should this be facilitated?

D. Cooperation of research infrastructures for curating processed (EBV) data products and publishing them in required formats

The research infrastructures contributing to the deliberations in the GLOBIS-B project have expressed their interest to cooperate at the global scale on interoperability to advance scientific progress in understanding and predicting the complexity of natural systems. By working together, the complementary data and capabilities of the cooperating infrastructures are thus better accelerated towards addressing the grand challenges on biodiversity and ecosystems. For example, the development and exploitation of cutting-edge IT technologies can contribute to data processing, supported by “machine-machine” interactions. Research infrastructures willing to cooperate in this regard are recommended to consider together the following:

1. Depending on policy guidance as outlined in the previous discussion items, joint efforts should be considered on developing better infrastructure support for EBV production with common support for running generic EBV computational workflows.
2. Become involved in the development and preservation of data, in access, integration and processing steps as an open-source and freely available service in a workflow-oriented e-infrastructure that supports curation, sharing and collaboration.
3. Promote standardized metadata and mechanisms to make EBV datasets ‘findable, accessible, interoperable and reusable’ (FAIR).

E. Overcoming legal constraints in relation to accessing data and achieving workflow interoperability

Constraints on legal interoperability emerge from restrictions on the use, modification and sharing of data and software. They have severe consequences for building EBV data products as they can impede the use of essential datasets and software tools, limit quality control, impede data aggregation and restrict re-usability. To overcome such constraints, datasets and software tools used for building EBV data products must be openly accessible and legally interoperable. More specifically, considerations for the research infrastructures are the following:

7. Legal interoperability can only be achieved if the legal use conditions are clearly and readily determined for each dataset and software tool, allowing for both creation and use of combined and/or derived products.

GLOBIS-B (654003)

7. Users should have the right to legally access and use datasets and software tools without seeking authorization from data right holders on a case-by-case basis.
7. EBV data products should be built from openly accessible data in open platforms, indicating for every dataset the legal use conditions.
7. As far as copyright or EU database protection apply to material used in EBV data products, they need to be expressly waived (by CC0-waiver) or mitigated (by CC-BY license or similar common-use-license).
7. EBVs, as well as underlying datasets, are developed in the public interest for scientific and policy purposes and for the sake of the conservation of biodiversity. They must be considered as belonging to the public domain.

Background information

GEO BON

The Group on Earth Observations Biodiversity Observation Network, or GEO BON, was established in 2008 as the biodiversity arm of GEO (Group on Earth Observations). Its mission is improving the acquisition, coordination, and delivery of biodiversity observations, and related ecosystem services, to both the scientific community, and decision-makers. GEO BON organizes its structure and its activities around two key components: the Essential Biodiversity Variables (EBVs), and the Biodiversity Observation Networks (BONs). GEO BON was approved as one of the four GEO Flagships at the GEO XIII Plenary in November 2016.

Essential Biodiversity Variables

The Essential Biodiversity Variables are a minimum set of measurements covering all dimensions of biodiversity, complementary to one another, that are required to study, report, and manage biodiversity change. The EBVs are produced by integrating primary biodiversity observations (e.g. in situ observations or remote-sensing products) and they constitute the building blocks of indicators of biodiversity change. The Essential Biodiversity Variables are organized in six EBV classes: Genetic Composition, Species Populations, Species Traits, Community Composition, Ecosystem Structure, and Ecosystem Function. Regardless of the class, these EBVs should be (1) able to capture critical scales and dimensions of biodiversity; (2) biological; (3) state variables (in general); (4) sensitive to change; (5) ecosystem agnostic (to the largest degree possible); and (6) technically feasible, economically viable and sustainable in time.

On this front, GEO BON and its partners have been developing a set of Global Indicators of Biodiversity Change that have been endorsed by the CBD to report on progress towards the Aichi Targets, as well as by the IPBES, in support of the Global Assessment. The indicators produced using the EBVs can also be used at the national and sub-national scales, for instance for National Biodiversity Strategies and Action Plans (NBSAPs), Environmental Impact Assessments, land and sea use planning, and biodiversity offsets. The EBVs can be understood as the level of integration between the raw biodiversity observations (derived either from in-situ observations or remote sensing), and the high-level indicators that are needed by stakeholders for reporting and decision-making. Several pilot applications of the EBV framework have either occurred or are being planned at national and regional scales in order to develop and test guidance on how the EBVs can be applied at various scales for improving biodiversity observation systems.

The GLOBIS-B project

GLOBIS-B (GLOBal Infrastructures for Supporting Biodiversity research) is a global cooperation funded by the Horizon 2020 research and innovation framework programme of the European Commission⁶. The main

⁶ The GLOBIS-B project received funding from the European Union's Horizon 2020 research and innovation

GLOBIS-B (654003)

aim of GLOBIS-B is to bring together biodiversity scientists, global research infrastructure operators and legal interoperability experts to identify the research needs and infrastructure services underpinning the concept of EBVs. The project is facilitating the multi-lateral cooperation of biodiversity research infrastructures worldwide⁷ and identifies the required primary data, analysis tools, methodologies and legal and technical bottlenecks to develop an agenda for research and infrastructure development to compute EBVs. This requires development of standards, protocols and workflows that are 'self-documenting' and openly shared to allow the discovery and analysis of data across large spatial extents and different temporal resolutions. The interoperability of existing biodiversity research infrastructures is crucial for integrating the necessary biodiversity data to calculate EBVs, and to advance the ability to assess progress towards the Aichi targets for 2020 of the Convention on Biological Diversity (CBD).

The GLOBIS-B project addressed these challenges by bringing together scientific, technical, and legal experts with infrastructure operators in a couple workshops to study possible approaches for EBVs on species distributions and abundances, on traits, and on species interactions. The majority of currently available data derives from incidentally reported observations or from surveys with standardized protocols. Enormous complexity exists in integrating these heterogeneous, multi-source data sets across space, time, taxa and different sampling methods. Integration of such data into global EBV data products requires correcting biases introduced by imperfect detection and varying sampling effort, dealing with different spatial resolution and extents, harmonizing measurement units from different data sources or sampling methods, applying statistical tools and models for spatial inter- or extrapolation, and quantifying sources of uncertainty and errors in data and models. The GLOBIS-B workshops resulted in the identification of generic workflow steps with multiple sequential activities, for example the identification and aggregation of various raw data sources, data quality control, taxonomic name matching and statistical modeling of integrated data.

Figure 1 summarizes key workflow steps for (species distribution) EBVs as part of the information supply chain, conceptually positioned between at the left side raw data and at the right side outcomes as indicators for reporting biodiversity change to policy and management. Three successive groups of workflow steps deal with (a) gathering EBV-usable data sets, (b) the production of harmonized EBV-ready data sets, and (c) the interpolation or extrapolation of these latter data sets with further analysis tools to result in derived and modelled data products. An example of a simple data product is shown for an EBV with the dimensions of taxonomy (e.g. species), time and space.

⁷ Atlas of Living Australia
 Biodiversity Committee of the Chinese Academy of Sciences
 Brazilian Reference Centre on Environmental Information
 DataONE Data Observation Network for Earth
 Global Biodiversity Information Facility
 GEO Biodiversity Observation Network
 Germplasm Bank of Wild Species at Kunming Institute of Botany
 LifeWatch European Infrastructure for Biodiversity and Ecosystem Research
 USA National Ecological Observatory Network
 South African National Biodiversity Institute.

GLOBIS-B (654003)

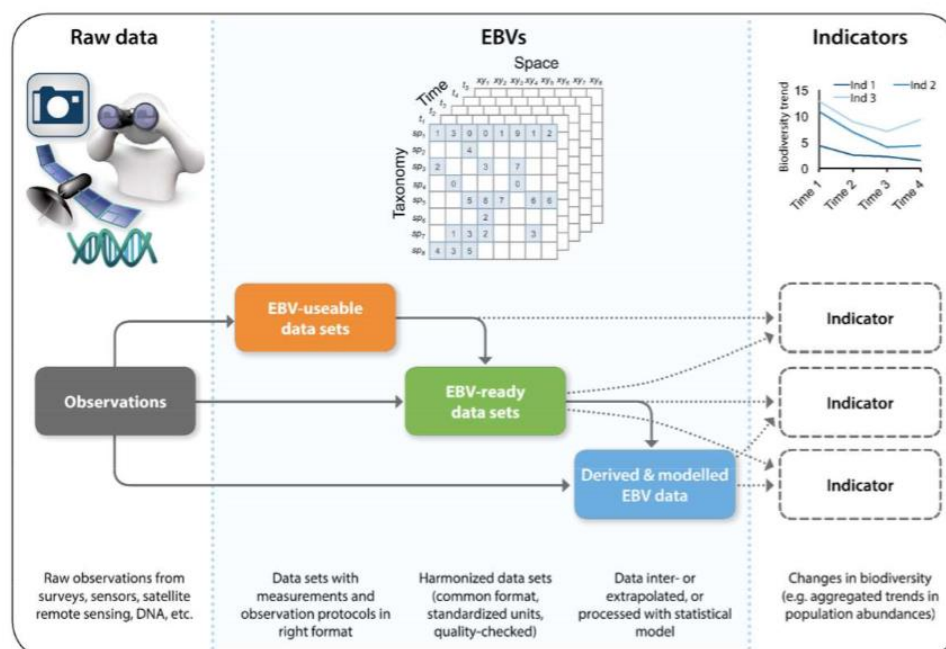


Fig. 2. Essential Biodiversity Variables as positioned between raw data and indicators.

Apart from scientific and technical challenges to implement and run such workflows as standardized method for producing EBVs for biodiversity indicators, also constraints related to legal interoperability must be solved. Such constraints emerge from restrictions on data use, modification and sharing, as well similarly for applied software. One of the most efficient approaches that help to assure legal interoperability are common-use licenses such as Creative Commons (CC) licenses. The ideal data sets for building EBV products are those in the public domain, with no restrictions on re-use and attribution (i.e. no need to specify source and license).

The identification of key workflow steps is highly relevant for building global EBV data products. This helps to establish analytical protocols that are robust, transparent and reusable, thereby improving reproducibility of ecological research. Existing projects, research infrastructures and citizen science efforts already operationalize many specific and generic workflow steps. Building reliable and representative global EBV data products requires filling of data gaps in geographic, temporal and taxonomic coverage. This necessitates a renewed effort in data mobilization and expanding existing biodiversity-monitoring initiatives worldwide, including citizen science projects. Producing EBVs in such way requires global cooperation based on common principles of best practices. To this end an international GLOBIS-B workshop in Bari drafted the following manifesto (February 2018).

The Bari Manifesto for Essential Biodiversity Variables (EBV) data products (summary)

(Prepared by a collective representation of the global biodiversity informatics research and data infrastructures community as an outcome of the GLOBIS-B project).

Exploiting primary data to produce and manage Essential Biodiversity Variables (EBV) data products depends on cooperation, practicality and interoperability among multiple stakeholders, including those

GLOBIS-B (654003)

collecting and mobilizing data with EBV potential (EBV-usable data), those making data 'EBV-ready' and those producing, publishing and preserving EBV data products. Ten principles offer best current practice guidance to data and infrastructure organizations to enhance their ability to contribute towards production of global EBV data products, whilst retaining autonomy and flexibility to achieve what is needed in ways appropriate to the organizations' own business. These principles are presented in the following "Bari Manifesto" as developed in an earlier GLOBIS-B workshop in Bari, Italy.

(1) Data Management Plan

Projects developing EBV data products should have comprehensive data management plans.

(2) Data Structure

EBV data products should adhere to agreed-upon minimum dimensions for each product (i.e., time, space, name/taxonomy (where applicable), etc.). All EBVs should be accommodated in a common framework that conforms (as far as is practically possible) with content and schema standards for representation formats and exchange protocols.

(3) Metadata

EBV data products and the EBV-ready datasets from which they are generated should have associated human- and machine-readable metadata, compliant with accepted community standards and sufficient for purposes of data discovery, access, fitness-for purpose evaluation, citation, interpretation and use.

(4) Services

EBV data products, EBV-ready datasets, digital objects and other related services should expose their capabilities and be accessible through common, standardized Application Programming Interfaces (APIs).

(5) Data Quality (Fitness-for-use)

Each EBV data product and EBV-ready dataset should include data quality documentation, sufficient to identify fitness-for-use of the data for specific purposes.

(6) Workflows

It should be possible to execute published, standard workflows for preparing, publishing and preserving EBV data products and the EBV-ready datasets from which they are produced. Ideally, such workflows should be defined and represented in a non-proprietary manner.

(7) Provenance

It should be possible to trace the EBV production process from the product back to the primary data and to reproduce the process. Provenance information must be readable both by humans and by machines.

(8) Ontologies / Vocabularies

EBV data products and EBV-ready datasets should be described by standard, openly accessible and machine-readable key vocabulary terms and conceptual relations (ontologies) presented in a simple way to enable wide usage.

(9) Data Preservation

EBV data products and associated underlying data should be preserved with an associated persistent identifier in a community supported and trusted repository.

(10) Accessibility

EBV data products and EBV-ready datasets must be sensitive, timely and FAIR (Findable, Accessible, Interoperable and Reusable).

Bibliografía

- Abadal, E. (2013). Gold or green: the debate on Open access policies. *International Microbiology*, 16(3), 109-203. Recuperado de <http://revistes.iec.cat/index.php/IM/article/view/74112>.
- Allard, S. (2012). DataONE: Facilitating eScience through Collaboration. *Journal of eScience Librarianship*, 1(1), e1004. Recuperado de <http://dx.doi.org/10.7191/jeslib.2012.1004>.
- Alonso García, E. (2008). *Introduction to International Environmental Law: Handbook with Cases and Materiales for American Lawyers*. Madrid: Servicio de Publicaciones de la Universidad Rey Juan Carlos.
- (2014). *Deliverable 4: Legal Interoperability*. Madrid: s. e.
- (2016a). *Deliverable 4.1 Legal input discussion document for workshop 1 (GLOBIS-B)*. Recuperado de http://www.globis-b.eu/files/intranet/Deliverables/Final/D4_1%20final.pdf.
- (2016b). Interoperabilidad legal problemas jurídicos derivados de manejar de datos científicos como «big data» por servicios web e infraestructuras de la economía del conocimiento: el ejemplo de los datos sobre diversidad biológica. *Revista de Privacidad y Derecho Digital*, 3, 141-179.
- Livia, B., De Leo, F., Hardisty, A., Keuchkerian, S., Konijn, J. *et al.* (2014). *Flock together with CReTIVE-B: A roadmap of global research data infrastructures supporting biodiversity and ecosystem science*. Recuperado de <http://orca.cf.ac.uk/88149>.
- Anderson, C. (2008). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. Recuperado de <https://www.wired.com/2008/06/pb-theory>.
- Andronico, G. *et al.* (2011). e-Infrastructures for e-Science: A Global View. *Journal of Grid Computing*, 9(2), 155-184.

- Appleyard, B. (2016). *Licences vs. Waivers*. Trabajo presentado en Issues in the Legal Interoperability of Research Data. Recuperado de <http://www.scidatacon.org/2016/sessions/84/paper/171>.
- Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., Moorman, D. (2004). An International Framework to Promote Access to Data. *Science*, 303(5665), 1777-1778. Recuperado de <http://science.sciencemag.org/content/303/5665/1777.full>.
- Arzberger, P., Uhler, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K. *et al.* (2004). Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal*, 3, 135-152. Recuperado de https://www.jstage.jst.go.jp/article/dsj/3/0/3_0_135/_pdf.
- Atkins, D., Droegemeier, K. K., Feldman, S. I., García Molina, H., Klein, M. L., Messerschmitt, D. G., Messina, P. *et al.* (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue Ribbon Advisory Panel on Cyberinfrastructure*. Arlington: National Science Foundation. Recuperado de <https://www.nsf.gov/cise/sci/reports/atkins.pdf>.
- Atlas of Living Australia. ALA Infrastructure Implementation* (2016). Recuperado de <http://www.ala.org.au/wp-content/uploads/2017/01/ALA-Infrastructure-Implementation-overview-October-2016-final.pdf>.
- Aubertin, C., Boisvert, V. y Vivien, F. (1998). La construction sociale de la question de la biodiversité. *Nature Sciences Sociétés*, 6(2), 7-19. Recuperado de [http://dx.doi.org/10.1016/S1240-1307\(97\)89512-X](http://dx.doi.org/10.1016/S1240-1307(97)89512-X).
- Ayala Mojica, P. (s. f.). *Redes y Telecomunicaciones*. Diapositiva n.º 31. Disponible en: <http://es.slideshare.net/jonenkairos/conceptos-redes-1>.
- Bailey, J. C. (2007). *What Is Open access? Digital Scholarship*. Houston: Digital Scholarship. Recuperado de <http://digital-scholarship.org/cwb/WhatIsOA.htm>.
- (2010). *Transforming scholarly publishing through Open access: a bibliography*. Houston: Digital Scholarship. Recuperado de <http://digital-scholarship.org/tsp/transforming.pdf>.
- Balmford, A., Bennun, L., Brink, B., Cooper, D., Côté, I., Crane, P., Dobson, A. *et al.* (2005). The Convention on Biological Diversity's 2010 Target. *Science* 307(5707), 212-213. Recuperado de <http://science.sciencemag.org/content/307/5707/212>.

- Band, J. y Kono, M. (2001). The Database Protection Debate in the 106th Congress. *Ohio State Law Journal*, 62(869), 1-7. Recuperado de http://moritzlaw.osu.edu/students/groups/oslj/files/2012/03/62.2.band_.pdf.
- Bargellini, P. Simonetta, C., Desnos, Y. L., Greco, B., Guidetti, V., Marchetti, P. G. *et al.* (2013). *Big Data from Space: Event Report. Frascati: ESA*. Recuperado de http://old.esaconferencebureau.com/docs/default-source/13c10_docs/here.pdf?sfvrsn=0.
- Barrionuevo, L. (2009). Acceso Abierto. Open access. *glossariumBITri*. Recuperado de <http://glossarium.bitrum.unileon.es/Home/acceso-abierto>.
- Bartha, G. y Kocsis, S. (2011). Standardization of geographic data: the European INSPIRE Directive. *European Joournal of Geography* 2(2), 79-89. Recuperado de <http://www.eurogeographyjournal.eu/articles/6-Bartha.pdf>.
- Basset, A. y Loss, W. (2012). Biodiversity e-Science: LifeWatch, the European infrastructure on biodiversity and ecosystem research. *Plant Biosystems. An International Journal Dealing with all aspects of Plant Biology*, 146(4), 780-782. Recuperado de <http://dx.doi.org/10.1080/11263504.2012.740091>.
- Beagrie, N., Lavoie, B. y Woollard, M. (2009). *Keeping Research Data Safe 2. Data Survey: Selection Criteria*. Bristol: JISC, Recuperado de http://www.data-archive.ac.uk/media/1687/KRDS2_finalreport.pdf.
- Beck, E. y Cowan, C. (1996). *Spiral Dynamics*. Oxford: Blackwell.
- Beck, S., Borie, M., Chilvers, J., Esguerra, A., Heubach, K., Hulme, M. *et al.* (2014). Towards a Reflexive Turn in the Governance of Global Environmental Expertise: The Cases of the IPCC and the IPBES. *GALA*, 23(2), 80-87. Recuperado de <https://pdfs.semanticscholar.org/e7af/403f5a9e24d8ef6e260d7a1dfa5eebb792f7.pdf>.
- Benkler, Y. (2000). From consumer to users: shifting the deeper structures of regulation towards sustainable commons and user access. *Federal Communications Law Journal*, 52(3), art. 9, 561-579. Recuperado de <http://www.repository.law.indiana.edu/fclj/vol52/iss3/9>.
- Berlin Declaration on Open access to Knowledge in the Sciences and Humanities* (2003). Múnich: Max-Planck-Gesellschaft. Recuperado de <https://openaccess.mpg.de/Berlin-Declaration>.

- Bernal, I. y Román Molina, J. (2016). *Gestión, difusión en acceso abierto y preservación de datos de investigación* (2.^a ed.). Madrid: Oficina Técnica de Digital. CSIC. Recuperado de http://digital.csic.es/bitstream/10261/138273/1/Datos2016_Bernal.pdf.
- Bernard, L., Craglia, M., Gould, M. y Kuhn, W. (2005). Towards an SDI Research Agenda. En K. Fullerton (ed.): *11th EC GIS & GIS Workshop. ESDI: Setting the Framework. Abstracts Handbook* (pp. 147-151). Sardinia: 11th EC GI y GIS Workshop.
- Berners-Lee, T. (2000). *Weaving the Web, the original design and ultimate destiny of the world wide web*. Nueva York: Harper Collins.
- Bethesda Statement on Open access Publishing*. (2003). Recuperado de <http://legacy.earlham.edu/~peters/fos/bethesda.htm>.
- Bioinformática (s. f.). En Wikipedia. Recuperado de <https://es.wikipedia.org/wiki/Bioinform%C3%A1tica>.
- Bisby, F. A. (2000). The Quiet Revolution: Biodiversity Informatics and the Internet. *Science*, 289(5488), 2309-2312. Recuperado de <http://science.sciencemag.org/content/289/5488/2309.full>.
- Björk, B., Welling, P., Laakso, M., Majlender, P., Hedlund, T. y Guðnason, G. (2010). Open access to the Scientific Journal Literature: Situation 2009. *PLoS ONE*, 5(6), e11273. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0011273>.
- Björk, B., Laakso, M., Welling, P. y Paetau, P. (2013). Anatomy of Green Open access. *Journal of the Association for Information Science and Technology*, 65, 237-250. Recuperado de https://helda.helsinki.fi/bitstream/handle/10138/157659/Bj_rk_Laakso_Welling_Paetau_2014_Anatomy_of_Green_OA.pdf?sequence=3.
- Blanke, J. M. (2002). Vincent Van Gogh sweat of the brow and database protection. *American Business Law Journal*, 39, 645-682.
- Blomquist, R. F. (2002). Ratification resisted. Understanding America's response to the Convention of Biological Diversity 1989-2002. *Golden Gate University Review*, 34(4), 494-586. Recuperado de <http://digitalcommons.law.ggu.edu/cgi/viewcontent.cgi?article=1844&context=ggulrev>
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access* (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital

- Information Recuperado de
http://blueribbontaskforce.sdsc.edu/biblio/BRTF_Final_Report.pdf
- Boettiger, S. y Bennet, A. B. (2006). Bayh dole: if we knew then what we know now. *Nature Biotechnology*, 24(3), 320-323. Recuperado de
www.pipra.org/Resources/publications/Boettiger_Bennett_2005.pdf.
- Bolton, J. (2001). Should we take global governance seriously? *Chicago Journal on International Law*, 1(2), 205-221. Recuperado de
<http://chicagounbound.uchicago.edu/cjil/vol1/iss2/2>.
- Borgman, C. (2010). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge (Massachusetts): The MIT Press.
- Borie, M. y Hulme, M. (2015). Framing global biodiversity: IPBES between mother earth and ecosystem services. *Environmental Science and Policy*, 54(diciembre), 487-496. Recuperado de
<http://www.sciencedirect.com/science/article/pii/S1462901115001069>.
- Boston Library (s. f.). *What is Research Data?* Recuperado de
<https://www.bu.edu/datamanagement/background/whatisdata>.
- Bovemberg, J. A. (2000). Should companies set up databases in Europe? *Nature Biotechnology*, 18(9), 907.
- Bowser, A., Wiggins, A. y Stevenson, R. D. (2013). *Data Policies for Public Participation in Scientific Research: a Primer*. Recuperado de
<http://www.birds.cornell.edu/citscitoolkit/toolkit/policy/Bowser%20et%20al%202013%20Data%20Policy%20Guide.pdf>.
- Braunstein, Y. (2002). *Repercusiones económicas de la protección de las bases de datos en los países en desarrollo y países en transición*. Ginebra: Organización Mundial de la Propiedad Intelectual. Recuperado de
http://www.wipo.int/meetings/es/doc_details.jsp?doc_id=2096.
- Breschi, S. y Cusmano, L. (2003). Unveiling the texture of a European Research Area: Emergence of oligarchic networks under EU Framework Programmes. *International Journal of Technology Management*, 27(8). Recuperado de
<ftp://ftp.zew.de/pub/zew-docs/evaluationR&D/LCusmano.pdf>.
- Broussard, S. L. (2007). The Copyleft Movement: Creative Commons Licensing. *Communication Research Trend*, 26(3), 3-17. Recuperado de
http://cscs.scu.edu/trends/v26/v26_n3.pdf.

- Brown, C. (2010). Copyleft the disguised copyright: why legislative copyright reform is superior to copyleft *licences*. *University of Missouri Kansas City Law Review*, 78, 749-785. Recuperado de: <https://ssrn.com/abstract=1688402>.
- Brunée, J. (2004). The United States and International Environmental Law: Living with an Elephant. *European Journal of International Law*, 15(4), 617-649.
- Budapest Open Access Initiative (2002). Recuperado de <http://www.budapestopenaccessinitiative.org>.
- Butchart, S.H.M., Di Marco, M. & Watson, J.E.M. (2016) Formulating smart commitments on biodiversity: lessons from the Aichi targets. *Conservation Letters*, 9, 457-468. Recuperado de <https://onlinelibrary.wiley.com/doi/full/10.1111/conl.12278>
- Burk, D. y Cohen, J. (2001). Fair use infrastructure for rights management system. *Harvard Journal of Law & Technology*, 15(1), 41-83. Recuperado de <http://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=1516&context=facpub>.
- Calzolari, A., Valerio, A., Capone, F., Napolitano, M., Villa, M., Pricci, F. *et al.* (2014). The European Research Infrastructures of the ESFRI Roadmap in Biological and Medical Sciences: status and perspectives. *Annali dell'Istituto Superiore di Sanita*, 50(2), 178-185. Recuperado de <http://www.scielosp.org/pdf/aiss/v50n2/12.pdf>.
- Candela, L., Castelli, D. y Pagano, P. (2009). D4Science: an e-Infrastructure for Supporting Virtual. En M. Agosti, F. Esposito y C. Thanos (eds.). *Post-proceedings of the 5th Italian Research Conference on Digital Libraries - IRCDL 2009* (166-169). Padua: DELOS y IRCDL. Recuperado de <https://pdfs.semanticscholar.org/0c18/ea62ce71bf63fa6a7ffe7ad55bca56b558f9.pdf#page=176>.
- Canhos, V. P., Souza, S., Giovanni, R. y Canhos, D. A. L. (2004). Global Biodiversity Informatics: setting the «scene» for a «new world» of ecological modeling. *Biodiversity Informatics*, 1, 1-13. Recuperado de <https://journals.ku.edu/index.php/jbi/article/view/3/1>.
- Cardinale, P. J. (2007). Sui generis database protection: second thoughts in the European Union and what it meant for the United States. *Chicago-Kent Journal on Intellectual Property*, 6(2), 157-176. Recuperado de <http://scholarship.kentlaw.iit.edu/cgi/viewcontent.cgi?article=1054&context=ckjip>.
- Cardozo, F. J. (s. f.). *Sistemas de telecomunicaciones. Concepto de IP en las nuevas redes*

- Integradas*. Recuperado de:
<http://www.monografias.com/trabajos33/telecomunicaciones/telecomunicaciones3.shtml>.
- Carpenter, S., Mooneyb, H. A., Agardc, J., Capistranod, D., DeFries, R. S., Díaz, S., Dietzg, T. *et al.* (2009). Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. *Proceedings of the National Academy of Sciences*, 106(5), 1305-1312. Recuperado de
<http://www.pnas.org/content/106/5/1305.full.pdf>.
- Carroll, M. W. (2008). *Complying with the National Institutes of Health Public Access Policy: Copyright considerations and options*. s. l.: A Joint Sparc; Science Commons; Arl White Paper. Recuperado de <http://sparcopen.org/wp-content/uploads/2016/01/COMPLYING-WITH-THE-NATIONAL-INSTITUTES-OF-HEALTH-PUBLIC-ACCESS-POLICY.pdf>.
- (2011). Why full open access matters. *PLoS Biology* 9(11), e1001210. Recuperado de
<http://www.journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001210>.
- (2015). Sharing Research Data: and Intellectual Property Law: A Primer. *PLoS Biology* 13(8): e1002235. Recuperado de
<https://doi:10.1371/journal.pbio.1002235>.
- Chavan, V. y Ingwersen, P. (2009). Towards a data publishing framework for primary biodiversity data: challenges and potentials for the biodiversity informatics community. *BMC Bioinformatics*, 10(S14), S2. Recuperado de
<http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-S14-S2>.
- Chen, R. (2016). *Legal Interoperability of Scientific Data: Why We Need It Now!* Trabajo presentado en Center for International Earth Science Information Network (Columbia University). Recuperado de
<http://www.scidatacon.org/2016/sessions/84/paper/362>.
- Christian, E. (2005). Planning for the Global Earth Observation System of Systems (GEOSS). *Space Policy*, 21(2), 105-109. Recuperado de
<http://www.sciencedirect.com/science/article/pii/S0265964605000226>.
- Clearwater, A. (2010). The New Ontologies: the effect of Copyright Protection on Public Scientific Data Sharing Using Semantic Web Ontologies. *The John Marshall Review of Intellectual Property Law*, 10(182), 183-205. Recuperado de
<http://repository.jmls.edu/cgi/viewcontent.cgi?article=1229&context=ripl>.

- Código abierto (s. f.). En Wikipedia. Recuperado de https://es.wikipedia.org/wiki/C%C3%B3digo_abierto.
- Comisión Europea (s. f.). *Belmont Forum*. Recuperado de <https://ec.europa.eu/research/environment/index.cfm?pg=belmont>.
- (14 de septiembre de 2016). Modernisation of the EU copyright rules. *European Commission*. Recuperado de <https://ec.europa.eu/digital-single-market/en/modernisation-eu-copyright-rules>.
- (4 de diciembre de 2017). *Dedicated Information session for ESFRI RIs and ERICs on EOSC and long-term sustainability of Research Infrastructures*. Recuperado de <https://ec.europa.eu/research/infrastructures/index.cfm>.
- Committee on Data of the International Council for Science (s. f.). CODATA's Mission. *CODATA*. Recuperado de <http://www.codata.org/about-codata/our-mission>.
- Conner, L. C., Gill, R. A. y O'Connor, R. (2013). Connecting to the Data-Intensive Future of Scientific Research. Recuperado de: <https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1006&context=spacegrant>
- Constable, H., Guralnick, R., Wieczorek, J., Spencer, C., Peterson, A. T. (2010). VertNet: A New Model for Biodiversity Data Sharing. *PLoS Biology*, 8(2), e1000309. Recuperado de <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1000309>.
- Contreras, J. L. y Reichman, J. H. (2015). Sharing by design: Data and decentralized commons. *Science*, 350(6266), 1312-1314. Recuperado de <http://science.sciencemag.org/content/350/6266/1312>.
- Convention on Biological Diversity (s. f.). *Conference of the Parties*. Recuperado de <http://www.cbd.int/cop>.
- (s. f.). *Role*. Recuperado de <https://www.cbd.int/secretariat/role.shtml>.
- Cook, T. (2012). Exceptions and limitations in European Union Copyright Law. *Journal of Intellectual Property Rights*, 17(mayo), 243-245. Recuperado de [http://nopr.niscair.res.in/bitstream/123456789/14081/1/JIPR%2017\(3\)%20243-245.pdf](http://nopr.niscair.res.in/bitstream/123456789/14081/1/JIPR%2017(3)%20243-245.pdf).
- Cordeiro, J. L. (2015). *Conferencia Prospectiva y las Tecnologías del Futuro*. Recuperado de <https://www.youtube.com/watch?v=izbFkfjeV7s>.
- Costello, M. J. (2009). Motivating online publication of data. *BioScience*, 59(5), 418-427. Recuperado de

- https://www.researchgate.net/profile/Mark_Costello/publication/37987920_Motivating_Online_Publication_Of_Data/links/02e7e5265953bebe03000000.pdf?origin=publication_list&ev=srch_pub_xdl.
- Michener, W., Gahegan, M. y Bourne, P. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution*, 28(8), 454-461. Recuperado de https://www.researchgate.net/profile/Philip_Bourne/publication/237146698_Biodiversity_data_should_be_published_cited_and_peer_reviewed/links/0046351d5958ae501f000000.pdf.
- Craglia, M. y Annoni, A. (2006). INSPIRE: An innovative Approach to the Development of Spatial Data Infrastructures in Europe. En H. Onsrud (ed.): *Research and Theory in Advancing Spatial Data Infrastructure Concepts* (pp. 93-105). Redlands (California): ESRI Press. Recuperado de <https://ec.europa.eu/jrc/en/publication/articles-books/inspire-innovative-approach-development-spatial-data-infrastructures-europe-concepts>.
- Craig, I. D., Plume, A. M., McVeigh, M. E., Pringle, J., Amin, M. (2007). Do open access articles have greater citation impact?: A critical review of the literature. *Journal of Informetrics*, 1(3), 239-248. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1751157707000466>.
- Creative Commons (s. f.). *CC0 1.0 Universal. Public Domain Dedication*. Recuperado de <https://creativecommons.org/publicdomain/zero/1.0>.
- Creative Commons (s. f.). *History*. Recuperado de <https://creativecommons.org/about/history>.
- Culliton, B. J. (1988). Authorship, Data Ownership examined. *Science*, 242, 658. Recuperado de <http://depts.washington.edu/uwbri/PDF%20Files/Authorship,%20Data%20Ownership%20Examined.pdf>.
- DataONE (s. f.). *DataONE Organization*. Recuperado de <https://www.dataone.org/organization>
- (s. f.). *What is DataONE*. Recuperado de <https://www.dataone.org/what-dataone>.
- Dato (s. f.). En Wikipedia. Recuperado de <http://es.wikipedia.org/wiki>.
- Davenport, T. H. (1997). *Information Ecology: Mastering the Information and Knowledge Environment*. Nueva York: Oxford University Press.

- David, P. A. y Foray, D. (2003). Economic fundamentals of the knowledge society. *Policy Futures in Education. An e-Journal*, 1(1), 1-22. Recuperado de <http://www-siepr.stanford.edu/workp/swp02003.pdf>.
- Davison, M. (2007). Database protection: Lessons form Europe, Congress and WIPO. *Case Western Reserve Law Review*, 57(4), 829-854. Recuperado de <http://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=1692&context=caselrev>.
- Decherney, P. (2014). Fair Use Goes Global. *Critical Studies in Media Communication*, 31(2), 146-152. Recuperado de <http://dx.doi.org/10.1080/15295036.2014.921321>.
- Demchenko, Y., Zhao, Z., Grosso, P., Wibisono, A., De Laat, C. (2013). Addressing big data Challenges for Scientific Data Infrastructure. En *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings* (pp. 48-55). Taipei: IEEE. Recuperado de <http://ieeexplore.ieee.org/abstract/document/6427494>.
- Derclaye, E. (2005). Databases sui generis right: what is a substantial investment? A tentative definition. *International Review of Intellectual Property and Competition Law*, 36(1), 2-30. Recuperado de <http://eprints.nottingham.ac.uk/27698/1/iic%20derclaye%202005.pdf>.
- Désirée, M. (2002). The CBD – Key Characteristics and Implications for Implementation. *Reciel*, 11(1), 17-28. Recuperado de <http://onlinelibrary.wiley.com/doi/10.1111/1467-9388.00299/abstract>.
- Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM*, 56(12), 64-73. Recuperado de <http://cacm.acm.org/magazines/2013/12/169933-data-science-and-prediction/fulltext>.
- Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M., Ash, N. *et al.* (2015). The IPBES Conceptual Framework. Connecting nature and people. *Current Opinion in Environmental Sustainability*, 14, 1-16. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1877343>.
- Dietz, T., Ostrom, E. y Stern, P. (2003). The Struggle to Govern the Commons. *Science*, 302(5652), 1907-1912. Recuperado de <http://science.sciencemag.org/content/302/5652/1907.full>.
- Di Marco, M., Watson, J.E.M., Venter O., Possingham, H.P. (2016). Global biodiversity targets require both sufficiency and efficiency, *Conservation Letters*, 9, pp. 395-397, Recuperado de <https://onlinelibrary.wiley.com/doi/epdf/10.1111/conl.12299>

- Doldirina, C., Eisenstadt, A., Onsrud, H. y Uhler, P. (en imprenta). *Legal approaches for Open Access to Research Data*.
- Doloreux, D. (2002). What should we know about regional systems of innovation. *Technnnology in Society* 24(3), 243-263. Recuperado de [http://dx.doi.org/10.1016/S0160-791X\(02\)00007-6](http://dx.doi.org/10.1016/S0160-791X(02)00007-6).
- Dreyfuss, R. (2004). Protecting the Public Domain of Science: Has the Time for an experimental use defense arrived? *Arizona Law Review*, 46(457), 458-472. Recuperado de <http://arizonalawreview.org/pdf/46-3/46arizlrev457.pdf>.
- Dubois, G., Clerici, M., Peedell, S., Mayaux, P., Bartholome, E., Gregoire, J. M. (2010). A Digital Observatory for Protected Areas - DOPA, A GEO BON contribution to the monitoring of Africa biodiversity. *Proceedings of Map Africa*, 23-25. Recuperado de http://www.eurogeoss.eu/Documents/publications%20-%20papers/2010%20Dubois%20etal_Map%20Africa%202010_Paper_v13.pdf.
- Duke, L. y Tech. Rev. 0017 (2001). The future of database protection in US Copyright law. *iBRIEF. Copyrights & Trademarks*. Recuperado de <https://scholarship.law.duke.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1016&context=dltr>.
- Duša, A., Nelle, D., Stock, G. y Wagner, G. G. (eds.) (2014). *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*. Berlín: Scivero. Recuperado de https://edoc.bbaw.de/files/2327/2014_FacingTheFuture_edoc.pdf.
- Dusollier, S. (2004). The Master's Tools v. The Master's House: Creative Commons v. Copyright. *Columbia Journal of Law & the Arts*, 29(3), 271-293. Recuperado de <http://www.crid.be/pdf/public/5421.pdf>.
- Earthzine (s. f.). *The Group on Earth Observations*. Recuperado de <https://earthzine.org/geo-and-geoss-the-group-on-earth-observations-and-the-global-earth-observations-system-of-systems>.
- Eberle, M. (1999). March-in rights under the Bayh-Dole Act: Public access to Federally Funded Research. *Marquette Intellectual Property Law Review*, 3(1), artículo 5, 155-180. Recuperado de <http://scholarship.law.marquette.edu/iplr/vol3/iss1/5>.
- Edwards, J. L. (2004). Research and Societal Benefits of the Global Biodiversity Information Facility. *Bioscience*, 54(6), 495-486. Recuperado de <http://bioscience.oxfordjournals.org/content/54/6/485.full>.

- Lane, M. A. y Nielsen, E. S. (2000). Interoperability of Biodiversity Databases: Biodiversity Information on Every Desktop. *Science*, 289(5488), 2312-2314. Recuperado de <http://science.sciencemag.org/content/289/5488/2312.full>.
- Edwards, P. E., Jackson, S. J., Bowker, G. C. y Williams, R. (2009). Introduction: An Agenda for Infrastructure Studies. *Journal of the Association for Information Systems* 10(5), 364-374. Recuperado de <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1508&context=jais>.
- Edwards, P. N., Jackson, S. J., Bowker, G. G. y Knobel, C. P. (2007). *Understanding Infrastructure: Dynamics, Tensions, and Design*. Ann Arbor (Michigan): Deep Blue.
- Edwards, P. N. Jackson, S., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D. *et al.* (2012). *Knowledge infrastructures: Intellectual Framework and Research Challenges*. Ann Arbor (Michigan): Deep Blue. Recuperado de <http://knowledgeinfrastructures.org/>: University of Michigan School of Information.
- Eechoud, V. (s. f.). Cuenta de Twitter personal. Recuperado de <https://twitter.com/EechoudMv>.
- Egloff, W., Agostí, D., Kishor, P. y Miller, J. (2017). Copyright and the use of images as Biodiversity Data. *Research Ideas and Outcomes*, 3(e12502). Recuperado de <http://biorxiv.org/content/early/2016/11/11/087015>.
- Eisemberg, R. y Nelson, R. (2002). Public vs. Proprietary Science: A Fruitful Tension? *Academic Medicine* 77(12 Pt 2), 1392-1399.
- Eisenberg, R. S. (1996). Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research. *Virginia Law Journal*, 82(8), 1663-1727. Recuperado de <http://repository.law.umich.edu/cgi/viewcontent.cgi?article=2223&context=articles>.
- Elkin-Koren, N. (2005). What contracts cannot do: the limits of private ordering in facilitating a Creative Commons. *Fordham Law Review*, 74(2), 375-422.
- Elliot, R. (2005). Who owns scientific data? The impact of intellectual property rights on the scientific publication chain. *Learned Publishing*, 18(2), 91-94. Recuperado de <http://www.d.umn.edu/~pschoff/documents/ElliotR05WhoOwnsScientificData.pdf>.

- Enke, N., Thessen, A., Bach, K., Bendix, J., Seeger, B., Gemeinholzer, B. (2012). The user's view on biodiversity data sharing. Investigating facts of acceptance and requirements to realize a sustainable use of research data. *Ecological Informatics* 11(septiembre), 25-33. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1574954112000222>.
- Erdozain López, J. (16 de septiembre de 2016). Comentario a la nueva Directiva (UE) 2016/943, sobre Secretos Comerciales. *Noticias Jurídicas*. Recuperado de <http://noticias.juridicas.com/conocimiento/articulos-doctrinales/11313-comentario-a-la-nueva-directiva-ue-2016-943-sobre-secretos-comerciales>.
- Esanu, J. y Uhler, P. (eds.) (2003). *The role of scientific and technical data and information in the public domain*. Washington, DC: National Academies Press. Recuperado de <https://www.nap.edu/read/10785/chapter/1#ii>.
- Europa Press (8 de mayo de 2017). La Junta participa en la primera asamblea del proyecto europeo 'LifeWatch'. *Europa Press*. Recuperado de <http://www.europapress.es/esandalucia/sevilla/noticia-junta-participa-primeras-asamblea-proyecto-europeo-lifewatch-20170508184022.html>.
- Fair use (s. f.). En Wikipedia. Recuperado de https://es.wikipedia.org/wiki/Fair_use.
- Faith, D., Collen, B., Ariño, A. H., Koleff, P., Guinotte, J. G., Kerr, J. *et al.* (2013). Bridging biodiversity data gaps: recommendations to meet user's data needs. *Biodiversity Informatics*, 8(2), 41-58. Recuperado de <https://journals.ku.edu/index.php/jbi/article/view/4126/4203>.
- Feld, C. K., Sousa, J. P., Da Silva, P. M. y Dawson, T. P. (2010). Indicators for biodiversity and ecosystem services: towards an improved framework for ecosystems assessments. *Biodiversity and Conservation*, 19(10), 2895-2919. Recuperado de <https://link.springer.com/article/10.1007/s10531-010-9875-0>.
- Fenter, P. (2017). Frontiers position statement: Impact of EU copyright reform on open science and innovation [entrada de blog]. *Frontier's blog*. Recuperado de <https://blog.frontiersin.org/2017/01/11/frontiers-position-statement-impact-of-eu-copyright-reform-on-open-science-and-innovation>.
- Fernández Muñoz, R. (s. f.): *Historia de Internet*. Recuperado de https://previa.uclm.es/Profesorado/Ricardo/Internet/Curso_de_Internet/Historia.htm.
- Field, C., DeFries, R. y Foster, D. *et al.* (2006). Integrated science and education plan for the National Ecological Observatory Network. *National Ecological*

- Observatory Network*. Recuperado de https://www.neonscience.org/sites/default/files/document-files/ISEP_2006Oct23.pdf.
- Fienberg, S. R., Martin, M. E. y Straf, M. L. (1985). *Sharing Research Data*. Washington DC: National Academy Press.
- Fishbein, S. A. (1991). Ownership of research data. *Journal of the Association of American Medical Colleges*, 66(3). Recuperado de http://journals.lww.com/academicmedicine/abstract/1991/03000/ownership_of_research_data.2.aspx.
- Fonseca, G. y Benson, P. J. (2003). Biodiversity conservation demands open access. *PLoS Biology*, 1(2): e46. Recuperado de <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0000046>.
- Foray, D. (2004). Innovative models for public-domain production of and open access to scientific and technical data and information. Introductory Remarks by Session Chair. En J. M., Esanu y P. F. Uhler (eds.): *Open access and the public domain in digital data and information for science: proceedings of an international symposium*. Washington DC: National Academies Press.
- Force 11 (s. f.). *Guiding Principles for Findable, Accessible, Interoperable and Re-Usable Data Publishing Version B1.0*. Recuperado de <https://www.force11.org/node/6062>.
- (s. f.). *The FAIR Data Principles*. Recuperado de <https://www.force11.org/group/fairgroup/fairprinciples>.
- Foss Lawyers Institute (s. f.). *Licencias sin copyleft*. Recuperado de http://esp.fossilawyers.org/?page_id=15.
- Foster, I. (2002). The Grid: a New Infrastructure for the XXIst Century Science. *Physics Today* 55(2), 27-42 Recuperado de <http://physicstoday.scitation.org/doi/full/10.1063/1.1461327?ver=pdfcov>.
- Free Software Foundation (2001). ¿Qué es el *software* libre? (trad. L. M. Arteaga Mejía). *El sistema Operativo GNU*. Recuperado de <http://www.gnu.org/philosophy/free-sw.es.html>.
- Frenzel, M., Klotz, S., Hardisty, A. y Banki, O. (2011). *LifeWatch. A European e-Science and observatory infrastructure supporting access and use of biodiversity and ecosystem data*. Recuperado de <http://proceedings.nature.com/documents/5843/version/1/files/npre20115843-1.pdf>.

- Gaijy, S., Chavan, V., Ariño, A. H., Otegui, J., Hobern, D., Sood, R. y Robles, E. (2013). Content assessment of the primary biodiversity data published through GBIF networks: status, challenges and potentials. *Biodiversity Informatics*, 8(2), 94-172. Recuperado de <https://journals.ku.edu/index.php/jbi/article/view/4124/4201>.
- Gaikwad, J. y Chavan, V. (2006). Open access and biodiversity conservation: challenges and potentials for the developping world. *Data Science Journal*, 5. Recuperado de https://www.jstage.jst.go.jp/article/dsj/5/0/5_0_1/_article.
- García Martínez, F. y Turégano, J. C. (2010). *Ciencias para el mundo contemporáneo. Guía de recursos didácticos. La revolución de la aldea global*. Canarias: Gobierno de Canarias.
- García Mexía, P. (2009). *Derecho Europeo de Internet, hacia la autonomía académica y la globalidad geográfica*. Madrid: NetBiblio.
- Gasser, U. y Bukert, H. (2003). Regulating Technological Innovation: An Information and a Business Law. En Rechtswissenschaftliche Abteilung der Universität St. Gallen (ed.): *Rechtliche Rahmenbedingungen des Wirtschaftsstandortes Schweiz* (pp. 503-523). Zürich: Dike.
- Gasser, U. y Palfrey, J. G. (2012). *Fostering Innovation and Trade in the Global Information Society: The Different Facets and Roles of Interoperability*. NCCR Trade Regulation Research Paper Series. Recuperado de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2192647.
- Gaston, K. J. y Spicer, J. I. (1998). *Biodiversity: an Introduction*. Oxford: Blackwell. Recuperado de http://birdi.ctu.edu.vn/birdi_cttt/noidung/giaotrich/BT313%20Biodiversity/Biodiversity%20-%20An%20Introduction.pdf.
- GBIF Secretariat Universitetsparken (s. f.). *What is GBIF*. Recuperado de <https://www.gbif.org/what-is-gbif#background>.
- GEO BON (s. f.). *What are EBVs*. Recuperado de <http://geobon.org/essential-biodiversity-variables/what-are-ebvs>.
- Gervais, D. (2008). Making Copyright Whole: A Principled Approach to Copyright Exceptions and Limitations. *University of Ottawa Law & Technology Journal*, 5(1 y 2), 1-41. Recuperado de <http://www.uoltj.ca/articles/vol5.1-2/2008.5.1-2.uoltj.Gervais.1-41.pdf>.

- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S. y Scott, P. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Londres: Sage.
- Giuliane, G., Ray, N., Schwarzer, S., De Bono, A., Peduzzi, P., Dao, H., Woerden, J., Witt, R. *et al.* (2013). Sharing Environmental Data through Geoss. En D. P. Albert y G. R. Dobbs (eds.): *Emerging methods and multidisciplinary applications in Geospatial Research* (pp. 266-281). Nueva York: Idea Group. Recuperado de <http://www.unige.ch/climate/Publications/Beniston/JAGR2011.pdf>.
- Glazer, D. *et al.* (2013). Data as IP and Data License Agreements. *Practical Law Company*. Recuperado de [https://content.next.westlaw.com/Document/I5f5951a21c8a11e38578f7ccc38dcbee/View/FullText.html?contextData=\(sc.Default\)&transitionType=Default&firstPage=true&bhcp=1](https://content.next.westlaw.com/Document/I5f5951a21c8a11e38578f7ccc38dcbee/View/FullText.html?contextData=(sc.Default)&transitionType=Default&firstPage=true&bhcp=1).
- GLOBIS-B (2015). Mission. Recuperado de <http://www.globis-b.eu/9/mission.html>. Goerg, C., Paulsch, A. y Neßhöver, N. (2010). A New Link Between Biodiversity Science and Policy. *GALA* 19(19), 183-186.
- Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, 8(4), 597-606. Recuperado de <http://nsuworks.nova.edu/cgi/viewcontent.cgi?article=1870&context=tqr>.
- Gómez-Díaz, T. (2015). *Software libre, software de código abierto, licencias*. Donde se propone un procedimiento de distribución de *software* y datos de investigación. *Hal Archives-ouvertes*. Recuperado de https://hal.archives-ouvertes.fr/hal-01261064/file/2015septTGD_FS_OSS_Lic_Proc_esp.pdf.
- González Aranda, J. M. (2011). Una aproximación a la ESFRI-LIFEWATCH: e-Science and Technology Infrastructure for Biodiversity Data and Observatories. En Consejería de Medioambiente de la Junta de Andalucía (org.): *III Jornadas sobre Información de Biodiversidad y Administraciones*. Celebradas en Sevilla el 23 de noviembre de 2011. Recuperado de <http://docplayer.es/7299852-Una-aproximacion-a-la-esfri-lifewatch-e-science-and-technology-infrastructure-for-biodiversity-data-and-observatories.html>.
- Rodríguez Clemente, R. y Lozano, S. (2009). e-Research in Research Cooperation Networks in Science & Technology Research. En M. Anandarajan y M. Anandarajan (eds.): *e-Research Collaboration* (pp. 167-199). Berlín; Heidelberg: Springer. Recuperado de https://link.springer.com/chapter/10.1007%2F978-3-642-12257-6_11.

- Gore, A. (1998). The Digital Earth: Understanding our planet in the 21st Century. *The Australian Surveyor*, 43(2), 89-91. Recuperado de <http://www.tandfonline.com/doi/abs/10.1080/00050348.1998.10558728>.
- Granjou, C., Mauz, I., Louvel, S. y Tournay, V. (2013). Assessing Nature? The Genesis of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES). *Science Technology & Society*, 18(1), 9-27. Recuperado de <http://journals.sagepub.com/doi/abs/10.1177/0971721813484232>.
- Griffith, S. (2005). Funding for Digital Libraries Research. *D-Lib Magazine*, 11(7 y 8). Recuperado de <http://www.dlib.org/dlib/july05/griffin/07griffin.html>.
- Grosheide, F. W. (2002). Data protection - The European Way. *Washington University Journal of Law and Policy*, 8(39). Recuperado de http://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1561&context=law_journal_law_policy.
- Group on Earth Observation (s. f.). *Data sharing and data management*. Recuperado de <http://www.earthobservations.org/dswg.php>.
- (s. f.). En Wikipedia. Recuperado de http://en.wikipedia.org/wiki/Group_on_Earth_Observations.
- Gunia, B. y Sandusky, R. B. (2010). Designing metadata for long-term data preservation: DataONE case study. *Proceedings of the Association for Information Science and Technology*, 47(1), 1-2. Recuperado de <http://onlinelibrary.wiley.com/doi/10.1002/meet.14504701435/full>.
- Guralnick, R. y Hill, A. (2009). Biodiversity informatics: automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics*, 25(4), 421-428. Recuperado de <http://bioinformatics.oxfordjournals.org/content/25/4/421.full>.
- Gurstein, M. (2011). Empowering the empowered or the effective data use for everyone? *First Monday*, 16(2). Recuperado de <http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764>.
- Guston, D. H. (2001). Boundary organizations in environmental policy and science: An introduction. *Science Technology & Human Values*, 26(4) [número especial: *Boundary Organizations in Environmental Policy and Science*], 399-408. Recuperado de http://www.atmos.washington.edu/~breth/PCC/guston_2001_BoundaryOrg.pdf.

- Halffter, G. (1994). ¿Qué es la biodiversidad? *Butlletí de la Institució Catalana d'Història Natural*, 62, 5-14. Recuperado de <http://publicacions.iec.cat/Front/repository/pdf/00000120%5C00000009.pdf>.
- Halpern, D. (2015) Data Management Principles Implementation Guidelines. GEO. 2015. Document 10. Recuperado de: <http://slideplayer.com/slide/9007507/>
- Hand, D., Mannila, H. y Smyth, P. (2001). *Principles of Data Mining*. Cambridge (Massachusetts): The MIT Press.
- Hannigan, J. (2014). *Environmental Sociology* (3.^a ed). Londres; Nueva York: Routledge.
- Hansen, H. S., Schroder, L., Hvingel, L. y Christiansen, J. S. (2011). Towards Spatially Enabled e-governance. A Case Study on SDI implementation. *International Journal of Spatial Data Infrastructures Research*, 6, 73-96. Recuperado de https://www.researchgate.net/publication/270891026_Towards_Spatially_Enabled_e-Governance_-_A_Case_Study_on_SDI_implementation.
- Hansen, M. C. y Loveland, T. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing Environment*, 122, 66-74. Recuperado de http://ac.els-cdn.com/S0034425712000314/1-s2.0-S0034425712000314-main.pdf?_tid=97320420-3d3c-11e7-b67e-00000aabb0f26&acdnat=1495271790_d87d6b211d7ae75d61d35113284b3ea8.
- Hardisty, A. (junio de 2015). Global Research Infrastructures for Biodiversity and Ecosystems Research [entrada de blog]. *Alex Hardisty*. Recuperado de <https://alexhardisty.wordpress.com/2015/06/17/global-research-infrastructures-for-biodiversity-and-ecosystems-research>.
- y Manset, D. (2014). CReATIVE-B *Deliverable D3.2: Guidelines for interoperability for biodiversity and ecosystem research infrastructures* [informe técnico]. Cardiff: Cardiff University, School of Computer Science and Informatics. Recuperado de <http://orca.cf.ac.uk/92562>.
- Roberts, D. y The Biodiversity Informatics Community (2013). A decadal view of biodiversity informatics: challenges and priorities. *BMC Ecology*, 13(16), 1-23. Recuperado de <https://bmcecol.biomedcentral.com/articles/10.1186/1472-6785-13-16>.
- Harison, E. (2010). Who owns enterprise information? Data ownership rights in Europe and the US. *Information & Management*, 47(2), 102-108. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0378720609001438>.

- Harmon, A. (21 de abril de 2010). Indian Tribe Wins Fight to Limit Research of its DNA. *New York Times*. Recuperado de <http://www.nytimes.com/2010/04/22/us/22dna.html>.
- Harnard, S. (2006a). Open access: the evidence and the verdict. *Journal of the Royal Society of Medicine*, 99. Recuperado de <http://eprints.soton.ac.uk/262894/1/KA06-18.PDF>.
- (2006b). Self-archiving should be mandatory. *Research Information*. Recuperado de <http://eprints.soton.ac.uk/262738/1/researchinf.html>.
- Harnard, S., Brody, T., Villiares, F., Carr, L., Hitchcock, S., Gingras, Y. *et al.* (2004). The Access/Impact Problem and the Green and Gold Roads to Open access. *Serials Review*, 30(4), 310-314. Recuperado de <http://www.tandfonline.com/doi/abs/10.1080/00987913.2004.10764930>.
- Harper, J. y Hawksworth, D. (1995). *Biodiversity: Measurement and Estimation*. Londres: Chapman and Hall.
- Harrop, S. H. y Pritchard, D. J. (2011). A hard instrument goes soft: The implications of the Convention on Biological Diversity's current trajectory. *Global Environmental Change*, 21(2), 474-480. Recuperado de <http://sro.sussex.ac.uk/44804>.
- Harvey F., Kuhn W., Pundt H., Bishr Y., and Riedemann C. 1999. "Semantic Interoperability: A Central Issue for Sharing Geographic Information." *The Annals of Regional Science* 33 (2): 213–232. doi: 10.1007/s001680050102
- Hernández Ernst, V., Poigné, A., Giddy, J., Hardisty, A., Voss, A. y Voss, H. (2009). Towards a Reference Model for the LifeWatch ICT Infraestructure. En *Focus on Life: contribuciones a la 39.ª Reunión Anual de la Sociedad de Ciencias de la Computación* (pp. 654-668). Reunión celebrada de 28 de septiembre al 2 de octubre de 2009 en Lübeck. Recuperado de <http://cs.emis.de/LNI/Proceedings/Proceedings154/gi-proc-154-15.pdf>.
- Hey, T. y Trefethen, A. (2002). The UK e-Science Core Programme and the Grid. *Future Generation Computer Systems*, 18(2012), 1017-1031. Recuperado de <http://users.ecs.soton.ac.uk/ajgh/FGCSPaper.pdf>.
- Tansley, S. y Tolle, K. (2009). *The Fourth Paradigm: Data-intensive Scientific Discovery*. Washington DC: Microsoft Research. Recuperado de <http://www.e-booksdirectory.com/details.php?ebook=3368>.
- Heywood, V. y Watson, R. T. (1995). *The Global Biodiversity Assessment*. Cambridge: Cambridge University Press.

- Hobern, D. (2003). *An Introduction to GBIF Biodiversity Informatics*. Copenhagen: DADI Programme Officer GBIF Secretariat.
- Apostolico, A., Arnaud, E., Bello, J. C., Canhos, D., Dubois, G. *et al.* (2012). *Global Biodiversity Informatics Outlook: Delivering biodiversity knowledge in the information age*. Copenhagen: Global Biodiversity Information Facility. Recuperado de <http://www.gbif.org/resource/80859>.
- Hodson, S. (2015). *Policies for open access to research data: international and scientific community dimension*. Recuperado de <http://helios-eie.ekt.gr/EIE/bitstream/10442/14566/1/HODSON.pdf>.
- Hoeren, T. (2014). Big data and the Ownership in Data: Recent Developments in Europe. *European Intellectual Property Review*, 12, 751-754. Recuperado de http://www.uni-muenster.de/Jura.itm/hoeren/veroeffentlichungen/Big_Data_and_the_Ownership_in_Data.pdf.
- Hoffmann, A., Penner, J., Vohland, K., Cramer, W., Doubleday, R., Henle, K. *et al.* (2014). Improved access to integrated biodiversity data for science, practice, and policy. The European Biodiversity Observation Network (EU BON). *Nature Conservation* 6, 49-65. Recuperado de <http://natureconservation.pensoft.net/articles.php?id=1349>.
- Hollingsworth, A., Uppala, S., Klinker, E., Burridge, D., Vitart, F., Onvlee, J. *et al.* (2005). The transformation of earth-system observations into information of socio-economic value in GEOSS. *Quarterly Journal of the Royal Meteorology Society*, 131(613), 3493-3512. Recuperado de <http://lw3ne3md9a.scholar.serialssolutions.com/?sid=google&auinit=A&au last=Hollingsworth&atitle=The+transformation+of+earth%E2%80%90system+observations+into+information+of+socio%E2%80%90economic+value+in+GEOSS&id=doi:10.1256/qj.05.181&title=Quarterly+jour>.
- Horn, J. (productor) y Akbarzadeh, A. (dir.). (2014). *Killswitch* [película]. Estados Unidos: Akorn Entertainment.
- Houghton, J. y Swan, A. (2013). Planting the seeds for a golden harvest: comments and clarifications on «going for gold». *D-Lib Magazine*, 19(1 y 2). Recuperado de <http://www.dlib.org/dlib/january13/houghton/01houghton.html>.
- Howard, A. (26 de julio de 2017). Notes from a roundtable on open data at eh White House [entrada de blog]. *Sunlight Foundation*. Recuperado de <https://sunlightfoundation.com/2017/07/26/white-house-open-data-roundtable>.

- Hrabanski, M. y Pesche, D. (2015). From the GBA to the IPBES: A comparative analysis of science-policy. En *International Conference on Public Policy* (pp. 1-21). Milán: ICPP. Recuperado de <http://agritrop.cirad.fr/579042>.
- Hugenholtz, B. (2001). *The new database right: early case from Europe*. Trabajo presentado en la Ninth Annual Conference on International IP Law & Policy, celebrada en Nueva York del 19 al 20 de abril de 2001. Recuperado de <http://www.ivir.nl/publicaties/download/fordham2001.pdf>.
- Hugenholtz, P. P. y Senftleben, M. (2011). *Fair use in Europe. In search of flexibilities*. Recuperado de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1959554: s.n.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446-456. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0740624X11000517>.
- Jasanoff, S. y Martello, M. (2004). *Earthly Politics: Local and Global. Environmental Governance*. Cambridge (Massachusetts): The MIT Press.
- Jeffries, M. (1997). *Biodiversity and conservation*. Londres: Routledge.
- Johnson, N. F. (2007). Biodiversity Informatics. *Annual Review of Entomology*, 52, 421-438. Recuperado de <http://www.annualreviews.org/doi/pdf/10.1146/annurev.ento.52.110405.091259>.
- Jones, M. B., Schildhauer, M. P., Reichman, O. J. y Bowers, S. (2006). The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere. *Annual Review of Ecology*, 37(2006), 519-544. Recuperado de <http://www.jstor.org/stable/pdf/30033842.pdf>.
- Jones, R. H. (1986). Is there a Property Interest in Scientific Research Data? *Berkeley Technology Law Journal*, 1(2), 447-482. Recuperado de <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1051&context=btlj>.
- Joppe, M. (2000). The Research Process. *The Quantitative Report Journal*, 8(4), 597-607.
- Kampe, T. U., Johnson, B. R., Kuester, M. y Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing*, 4(1), 043510. Recuperado de <http://remotesensing.spiedigitallibrary.org/article.aspx?articleid=707925%20>

- Kantardzic, M. (2011). *Data Mining: Concepts, Models, Methods, and Algorithms* (2.^a ed.). Canadá: IEEE Press; Wiley.
- Karlsson, S. (2007). The North-South Knowledge Divide: Consequences for Global Environmental Governance. En D. C. Esty y M. H. Ivanova (eds.): *Global Environmental Governance: Options and Opportunities* (pp. 53-76). Connecticut: Yale School of Forestry & Environmental Studies. Recuperado de <http://environment.yale.edu/publication-series/documents/downloads/h-n/karlsson.pdf>.
- Keller, M., Schimel, D. S., Hargrove, W. W. y Hoffman, F. M. (2008). A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment*, 6(5), 282-284. Recuperado de https://www.climatemodeling.org/~forrest/pubs/papers/Keller_FrontEcolEnviron_20080601.pdf.
- Keller, P. (9 de agosto de 2016). Copyright Reform: European Commission does the rightsholders' bidding [entrada de blog]. *Communia*. Recuperado de <http://www.communia-association.org/2016/08/29/copyright-reform-european-commission-rightsholders-bidding>.
- Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G. *et al.* (2009). Data-intensive Science: A New Paradigm for Biodiversity Studies. *BioScience*, 59(7), 613-620. Recuperado de <https://academic.oup.com/bioscience/article/59/7/613/334937/Data-intensive-Science-A-New-Paradigm-for>.
- Kirk, J. y Miller, M. L. (1986). *Reliability and validity in qualitative research*. California: Sage.
- Kissling, D., Ahumada, J.A., Bowser, A., Fernandez, M.J., Fernandez, N., Alonso García, E., Guralnick, R.P., Isaac, N.J.B., Kelling, S., Los, W., McRae, L., Mihoub, J.-B., Obst, M., Santamaría, M., Skidmore, A.K., Williams, K.J., Agosti, D., Amariles, D., Arvanitidis, C., Bastin, L., De Leo, F., Egloff, W., Elith, J., Hobern, D., Martin, D., Pereira, H.M., Pesole, G., Peterseil, J., Saarenmaa, H., Schigel, D., Schmeller, D.S., Segata, N., Turak, E., Uhler, P., Wee, B., Hardisty, A.R. (2017) Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev.* 2017. Recuperado de <https://onlinelibrary.wiley.com/doi/full/10.1111/brv.12359>.
- Kissling, D., Hardisty, A., Alonso García, E. y Santamaría, M., De Leo, F., Pesole, G. *et al.* (2015). Towards global interoperability for supporting biodiversity

- research on essential biodiversity variables (EBVs). *Biodiversity*, 16(2 y 3), 99-107. Recuperado de <http://dx.doi.org/10.1080/14888386.2015.1068709>.
- Koetz, T. F. K. y Bridgewater, P. (2012). Building better science-policy interfaces for international environmental governance: assessing potential within the Intergovernmental Platform for Biodiversity and Ecosystem Services. *International Environmental Agreements: Politics, Law and Economics*, 12(1), 1-21. Recuperado de <https://link.springer.com/article/10.1007/s10784-011-9152-z>.
- Koezt, T., Bridgewater, P., Van den Hove, S. y Siebenhüner, B. (2008). The role of the Subsidiary Body on Scientific, Technical and Technological Advice to the Convention on Biological Diversity as science policy interface. *Environmental Science & Policy*, 11(6), 505-516. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1462901108000464>.
- Konijn, J. (2016). *D2.2 Report of Workshop 1*. Recuperado de http://www.globis-b.eu/files/intranet/Deliverables/Final/Deliverable%202_2%20v2_0%20final.pdf.
- Kormos, C., Grosko, B. y Mittermeier, R. (2001). U.S. participation in international environmental law and policy. *Georgetown International Environmental Law Review*, 13, 661-693.
- Korn, N., Oppenheim, C. y Duncan, C. (2007). *IPR and Licensing issues in Derived Data*. Recuperado de <http://webarchive.nationalarchives.gov.uk/20140702233839/http://www.jisc.ac.uk/media/documents/projects/iprinderiveddatareport.pdf>.
- Kroes, N. (s. f.). *Digital Agenda for Europe*. Recueperado de https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/dae_v6.pdf.
- Kvalheim, V. y Kvamme, T. (2014). Policies for Sharing Research Data in Social Sciences and Humanities: A survey about research funders' data policies. *International Federation of Data Organisations*. Recuperado de <http://ifdo.org/wordpress/ifdo-report-2014-policies-for-sharing-research-data-in-social-sciences-and-humanities>.
- Landsbergen, D. y Wolken, D. (1998). *Eliminating legal and policy barriers to interoperable government systems*. Ohio: Ohio Supercomputer Center. Recuperado de https://www.osc.edu/files/press/releases/1998/phase_2_Recommendations.pdf: Ohio Supercomputer Center, ECLIPS PROGRAM.
- Larigauderie, A. (2015). *Diversitas*.

- y Mooney, H. (2010). The Intergovernmental science-policy Platform on Biodiversity and Ecosystem Services: moving a step closer to an IPCC-like mechanism for biodiversity. *Current Opinion in Environmental Sustainability*, 2(1 y 2), 9-14. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1877343510000072>.
- Latenbacher, C. L. (2006). The Global Earth Observation System of Systems: Science Serving Society. *Space Policy* 22(1), 8-11. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0265964605001220>.
- Laurentis, D. de, Dickerson, C., DiMario, M., Gartz, P., Jamshidi, M. M., Nahavandi, S. *et al.* (2007). A Case for an International Consortium on System-of-Systems Engineering. *IEEE Systems Journals*, 1(1), 68-71. Recuperado de <https://dro.deakin.edu.au/eserv/DU:30007797/nahavandi-caseforaninternational-2007.pdf>.
- Le Prestre, P. (2002a). *The Convention on biological diversity and the construction of a new biological order*. Brookfield. Ashgate.
- (2002b). The operation of the CBD Convention Governance System. En P. Le Prestre (ed.): *Governing global biodiversity: The evolution and implementation of the convention on biological diversity* (pp. 91-114). Aldershot: Ashgate.
- (ed.) (2002c). *Governing global biodiversity: The evolution and implementation of the convention on biological diversity*. Aldershot: Ashgate.
- (2008) The CBD at ten: The long road to effectiveness, *Journal of International Wildlife Law & Policy*, 5:3, (pp269-285), DOI: 10.1080/13880290209354014.
- y Compagnon, D. (2017). IPBES and governance of the international biodiversity regime complex. En M. Hrabanski y D. Pesche (eds.): *The Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES)* (pp. 18-39). Nueva York: Routledge.
- Lee, P. (2008). The Evolution of intellectual infrastructure. *Washington Law Review*, 83(3); *UC Davis Legal Studies Research Paper Series*, 137. Recuperado de https://lw3ne3md9a.scholar.serialssolutions.com/?sid=google&auinit=P&a_ultast=Lee&atitle=The+Evolution+of+Intellectual+Infrastructure&title=Washington+law+review&volume=83&date=2008&spage=39&issn=0043-0617.
- Lessig, L. (2001). *The future of ideas*. Nueva York: Random House. Recuperado de http://www.the-future-of-ideas.com/download/lessig_FOI.pdf.

- (2002). The architecture of innovation. *Duke Law Journal*, 51(6), 1783-1801. Recuperado de <https://scholarship.law.duke.edu/dlj/vol51/iss6/2>.
- (2004a). *Free Culture: how big media uses technology and the law to lock down culture and control creativity*. Nueva York: The Penguin Press. Recuperado de <http://www.free-culture.cc/freeculture.pdf>.
- (2004b). The Creative Commons. *Montana Law Review*, 65, 1-13.
- Leval, P. (1990). Toward a fair use standard. *Harvard Law Review*, 103, 1105-1136. Recuperado de https://www.law.berkeley.edu/files/Leval_-_Fair_Use.pdf.
- LifeWatch (2015). *Scientific and Technical description of LifeWatch Eric*. Recuperado de <https://www.lifewatch.eu/documents/20181/21696/Scientific+and+Technical+description+of+LifeWatch+ERIC/212169f2-3b02-4357-8272-66b55c978fdc>.
- López López, J. C. (27 de febrero de 2014). La moda del *big data*: ¿En qué consiste en realidad? *El Economista*. Recuperado de <http://www.economista.es/tecnologia/noticias/5578707/02/14/La-moda-del-Big-Data-En-que-consiste-en-realidad.html>.
- Loreau, M., Oteng-Yeboah, A., Arroyo, M. T. K., Babin, D., Barbault, R., Donoghue, M. *et al.* (2006). Diversity without representation. *Nature*, 442, 245-246. Recuperado de <https://www.nature.com/nature/journal/v442/n7100/full/442245a.html>.
- Lossau, N. (2012). An Overview of Research Infrastructures in Europe. And Recommendations to LIBER. *LIBER Quarterly*, 21(3-4), 313-329. Recuperado de <http://doi.org/10.18352/lq.8028>.
- Lovejoy, T. (1980). Changes in biological diversity. En B. Go (ed.): *The Global 2000 report to the President*, vol. 2: *The technical report* (pp. 327-332). Harmondsworth: Penguin Books.
- Lubens, R. (2003). Survey of developments in European Database Protection. *Berkeley Technology Law Journal* 18, p. 441 Recuperado de <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1417&context=btlj>.
- MacGregor, J. y Stranack, K. y Willinsky, J., (1998). *The Public Knowledge Project: Open Source Tools for Open access to Scholarly Communication*. Recuperado de https://link.springer.com/chapter/10.1007/978-3-319-00026-8_11.
- Madison, M. J. (2004). A Pattern-Oriented Approach to Fair Use. *William & Mary Law Review*, 54(4), 1525-1690. Recuperado de

<http://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=1333&context=wmlr>.

Malcolm, J. (15 de junio de 2015). European Parliamentarians to Vote Tomorrow: Fix Our Broken Copyright Law [entrada de blog]. *Electronic Frontier Foundation*. Recuperado de <https://www.eff.org/deeplinks/2015/06/european-parliamentarians-vote-tomorrow-fix-our-broken-copyright-law>.

Mansell, R. (2014). Here comes the Revolution. The European Digital Agenda. En K. Donders, C. Pauwels y J. Loisen (eds.): *The Palgrave Handbook of European Media Policy* (pp. 202-217). Londres: Palgrave Macmillan. Recuperado de https://www.academia.edu/7776435/Here_Comes_the_Revolution_The_European_Digital_Agenda.

Marques, A., Pereira, H. M., et al. (2014). A framework to identify enabling and urgent actions for the 2020 Aichi Targets. *Basic and Applied Ecology*. Volume 15, Issue 8, December 2014, Pages 633-638. Recuperado de <https://doi.org/10.1016/j.baae.2014.09.004>.

Martellos, S. y Attorre, F. (2012). New trends in biodiversity informatics. *Plant Biosystems*, 146(4), 749-751. Recuperado de <http://www.tandfonline.com/doi/abs/10.1080/11263504.2012.740092>.

Maurer, S. E. (2005). Across two worlds database protection in the United States and Europe. En J. Putnman (ed.): *Intellectual Property and Innovation in the Knowledge-Based Economy Conference Proceeding*. Calgary: University of Calgary Press. Conferencia celebrada el 23 y 24 de mayo de 2005 en Toronto (Canadá). Recuperado de https://www.researchgate.net/publication/228794091_Across_Two_Worlds_Database_Protection_in_the_United_States_and_Europe.

Maurer, S. M. y Scotchmer, S. (1999). Database protection: Is it broken and should we fix it? *Science*, 284(5417), 1129-1130. Recuperado de <http://search.proquest.com/docview/213562477?accountid=14475>.

Max Planck (s. f.). The Three-Step Test. *Electronic Frontier Foundation*. Recuperado de https://www.eff.org/files/filenode/three-step_test_fnl.pdf.

Mayer-Schonberger, V. y Cukier, K. (2013). *Big data: A Revolution That Will Transform how We Live, Work and Think*. Londres: John Murray.

Mayer, K. (s. f.). *From Science 2.0 to Open science: turning rethoric into action?* Recuperado de https://www.researchgate.net/profile/Katja_Mayer2/publication/27945002

- 5_From_Science_20_to_Open_Science_-_Turning_rhetoric_into_action/links/55931da308ae1e9cb4298c8d.pdf.
- Mayernik, M., Wallis, J. C., Pepe, A. y Borgman, C. L. (2008). *Whose data do you trust? Integrity issues in the preservation of scientific data*. Recuperado de https://www.ideals.illinois.edu/bitstream/handle/2142/15119/PA10-1_iconf08.pdf?sequence=2&isAllowed=y).
- McAfee, A. y Brynjolfsson, E. (2012). Big data: The Management Revolution, *Harvard Business Review*, 10(2012), 60-68. Recuperado de http://www.rosebt.com/uploads/8/1/8/1/8181762/big_data_the_management_revolution.pdf.
- McManis, C. R. (2001). Database Protection in the Digital Information Age. *Roger Williams University Law Review*, 7(1), 7-45. Recuperado de http://docs.rwu.edu/cgi/viewcontent.cgi?article=1215&context=rwu_LR.
- MEA (Millennium Ecosystem Assessment) (2005). *Ecosystems and human well-being: Synthesis*. Washington: Island Press.
- Mediano, S. (6 de abril de 2016). La directiva de Secreto Profesional será una realidad [entrada de blog]. *Santiago Mediano Abogados*. Recuperado de <http://www.santiagomediano.com/directiva-relativa-al-secreto-profesional-sera-una-realidad>.
- Meester, L. de, Tienderenb, P., Wergerc, M., Hectord, A., Wörheidee, G., Niemelä, J., Aguilar, A. *et al.* (2011). Challenges for biodiversity research in Europe. *Procedia Social and Behavioral Sciences* 13, 83-100.
- Merton, R. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago; Londres: The University of Chicago Press.
- Meyer, E. T., Schroeder, R. y Dutton, W. H. (2008). The Role of e-Infrastructures in the Transformation of Research Practices and Outcomes. *iConference 2008 Papers*. Recuperado de <https://www.ideals.illinois.edu/handle/2142/15120>.
- Michener, W. (21 de marzo de 2012). Five New Paradigms for Science and an Introduction to DataONE [entrada de blog]. *Educause Review*. Recuperado de <http://er.educause.edu/articles/2012/3/five-new-paradigms-for-science-and-an-introduction-to-dataone>.
- (2015). Ecological data sharing. *Ecological informatics*, 29(1), 33-44. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1574954115001004>.

- Michener, W. A., Allard, S., Budden, A., Cook, R. B., Douglass, K., Frame, M. *et al.* (2012). Participatory design of DataONE. Enabling cyberinfrastructure for the biological and environmental sciences. *Ecological Informatics*, 11(2012), 5-15. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1574954111000768>.
- Michener, W. K. y Jones, M. B. (2012). Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology & Evolution* 27(2), pp. 85-93. Recuperado de <https://doi.org/10.1016/j.tree.2011.11.016>.
- Michener, W., Vieglais, D., Vision, T., Kunze, J., Cruse, P. y Janée, G. (2011). DataONE: Data Observation Network for Earth. Preserving Data and Enabling Innovation in the Biological and Environmental Sciences. *D-Lib Magazine*, 17(1 y 2). Recuperado de <http://www.dlib.org/dlib/january11/michener/01michener.print.html>.
- Miller, C. (2001). Challenges in the application of Science to global affairs: Contingency, trust, and moral order. En P. N. Edwards y C. A. Miller (eds.): *Changing the atmosphere. Expert knowledge and environmental governance* (pp. 247-285). Boston: The MIT Press.
- Minster, J. B. y Clement, G. P. (2016). Data Sharing Entails Shared Responsibilities. Conference Paper. Parte de la sesión SCIDATACON 2016, presentada por Robert Chen. Recuperado de <http://www.scidatacon.org/2016/sessions/84/>.
- Mokyr, J. (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Nueva Jersey: Princeton University Press.
- Molloy, J. C. (2011). The Open Knowledge Foundation: Open data Means Better Science. *PLoS Biology* 9(12), e1001195. Recuperado de <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001195>.
- Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L. O. B. y Wilkinson, M. D. (2017). Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services & Use* 37(1), 49-56. Recuperado de <https://content.iospress.com/articles/information-services-and-use/isu824>.
- Montes, S. (2011). *Licencias FLOSS*. Recuperado de <http://es.slideshare.net/sergioslideshare/licencias-floss>.
- Moodley, D., Vahed, A., Simonis, I., McFerren, G., van Zyl, T., *et al.* (2008). Enabling a new era of Earth observation research: scientific workflows for

- the Sensor Web. *Ecological Circuits*, 1, 20-30. Recuperado de <http://www.ee.co.za/wp-content/uploads/legacy/8%20Scientific%20workflows%20for%20the%20Sensor%20Web.pdf>.
- Moore, K. (1996). Organizing Integrity: American Science and the creation of public interest organizations 1955-1975. *American Journal of Sociology*, 101(6), 1592-1627. Recuperado de http://www.jstor.org/stable/2782113?seq=1#page_scan_tab_contents.
- Morin, J. F., Louafi, S., Orsini, A. y Oubenal, M. (2016). Boundary Organizations in Regime Complexes: A Social Network Profile of IPBES. *Journal of International Relations and Development*, 20(3), 543-577. Recuperado de <https://link.springer.com/article/10.1057/s41268-016-0006-8>.
- Moritz, T., Krishnan, S., Roberts, D., Ingwersen, P., Agosti, D., Penev, L. et al. (2011). Towards mainstreaming of biodiversity data publishing: recommendations of the GBIF Data Publishing Framework Task Group. *BMC Bioinformatics*, 12(S15). Recuperado de <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-S15-S1>.
- Moshe, J. y Teubald, M. (1986). Innovation policy in an open economy: A normative framework for strategic and tactical issues. *Research Policy*, 15(3), 121-138. Recuperado de <http://www.sciencedirect.com/science/article/pii/0048733386900077>.
- Murray-Rust, P. (2008). Open data in Science. *Serials Reviews*, 34(1), 52-64. Recuperado de <http://precedings.nature.com/documents/1526/version/1>.
- Neylon, C., Pollock, R. y Wilbanks, J. (2010). *Panton Principles, Principles for open data in science*. Recuperado de <http://pantonprinciples.org>.
- Murray, F. y Stern, S. (2006). Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior & Organization*, 63(4), 648-687. Recuperado de <http://www.nber.org/papers/w11465>.
- Naciones Unidas (s. f.). *Convenio sobre la Diversidad Biológica*. Recuperado de <http://www.un.org/es/events/biodiversityday/convention.shtml>.
- Nativi, S., Craglia, M., Vaccari, L. y Santoro, M. (2011). Searching the New Grail: Inter-Disciplinary Interoperability. En S. Geertman, W. Reinhardt y F. Toppen (eds.): *Proceedings of the 14th AGILE International Conference on*

- Geographic Information Science*, celebrada del 18 al 22 de abril de 2011 en Utrecht (Países Bajos).
- Nativi, S., Mazzetti, P., Santoro, M., Papeschi, F., Craglia, M. y Ochiai, O. (2015). Big data challenges in building the Global Earth Observations System of Systems. *Environmental Modelling & Software*, 68, 1-26. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1364815215000481>.
- Navarro, M. (dir.), Magro, E., Lorenz, U., Parrilli, M. D., Karlsen, J. y Egurbide, I. (2013). *Las infraestructuras de conocimiento: el caso vasco desde una perspectiva internacional*. Bilbao: Instituto Vasco de Competitividad-Fundación Deusto. (Colección Cuadernos Orkestra, 3). Recuperado de <http://www.orkestra.deusto.es/es/investigacion/publicaciones/cuadernos-orkestra/299-infraestructuras-conocimiento-caso-vasco-perspectiva-internacional>.
- Nentwich, M. y König, R. (2003). *Cyberscience 2.0: Research in the Age of Digital Networks*. Viena: Austrian Academy of Sciences.
- Neon (s. f.). *The Observatory*. Recuperado de <http://www.neonscience.org/observatory>.
- Nesshöver, C., Vandewalle, M., Wittmer, H., Balian, E. V., Carmen, E., Geijzendorffer, I. R. et al. (2016). The Network of Knowledge approach: improving the science and society dialogue on biodiversity and ecosystem services in Europe. *Biodiversity and Conservation*, 25(7), 1215-1233. Recuperado de <http://link.springer.com/article/10.1007/s10531-016-1127-5>.
- Nicholson, E., Collen, B., Barausse, A., Blanchard, J. L., Costelloe, B. T., Sullivan, K. M. E. et al. (2012). Making robust policy decisions using global biodiversity indicators. *PLoS ONE*, 7(7): e41128, p. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0041128>.
- Nielsen, M. (2012). *Reinventing discovery: the new era of networked science*. Nueva Jersey: Princeton University Press.
- NIH Public Access Policy (s. f.). en Wikipedia. Recuperado de https://en.wikipedia.org/wiki/NIH_Public_Access_Policy.
- Norse, E. y McManus, R. (1980). Ecology and living resources biological diversity. En Council on Environmental Quality (ed.): *Environmental Quality. The eleventh Report of the Council On Environmental Quality* (pp. 31-80). Washington DC: s. e.
- Núñez, I., Barahona, A. y González-Gaudio, E. (2003). La Biodiversidad: historia y contexto de un concepto. *Interciencia*, 28(7), 387-393. Recuperado de

- http://datateca.unad.edu.co/contenidos/201602/2014-2_201602/Concepto_de_biodiversidad.pdf.
- O'Brien, T. (20 marzo 2017). Trump's quiet war on data begins [entrada de blog]. *Engadget*. Recuperado de <https://www.engadget.com/2017/03/20/trumps-quiet-war-on-data-begins>.
- O'Brien, W. (2013). March-in Rights Under the Bayh-Dole Act: The NIH's Paper Tiger? *Seton Hall Law Review*, 43(2013), 1043-1432. Recuperado de <http://scholarship.shu.edu/cgi/viewcontent.cgi?article=1488&context=shlr>.
- O'Connor, B., C. Secades, J. Penner, R. Sonnenschein, A. Skidmore, N. D. Burgess, et al. (2015). Earth observation as a tool for tracking progress towards the Aichi biodiversity targets. *Remote Sens. Ecol. Conserv.* 1, 19–28. Recuperado de <https://zslpublications.onlinelibrary.wiley.com/doi/epdf/10.1002/rse2.4>.
- OECD (1999). *Megascience Forum Working Group on Biological Informatics*. Recuperado de <http://www.oecd.org/science/sci-tech/2105199.pdf>.
- (2015). *OECD Digital Economy Outlook 2015*. París: OECD Publishing.
- Ogburn, J. L. (2016). Extending the principles and promise of scholarly communication reform. A Chronicle and future glimpse. En K. L. Smith y K. A. Dickson (eds.): *Open Access and the Future of Scholarly Communication: Policy and Infrastructure* (pp. 1-30). Londres: Rowman & Littlefield.
- Oliver, T. H., Heard, M. S., Isaac, N. J., Roy, D. B., Procter, D., Eigenbrod, F. et al. (2015). Biodiversity and resilience of ecosystem functions. *Trends Ecology*, 30(11), 673-684. Recuperado de [http://www.cell.com/trends/ecology-evolution/abstract/S0169-5347\(15\)00218-9](http://www.cell.com/trends/ecology-evolution/abstract/S0169-5347(15)00218-9).
- Onsrud, H. J. y López, X. (1998). Intellectual Property Rights in Disseminating Digital Geographic Data, Products, and Services: Conflicts and Commonalities among European Union and United States Approaches. En I. Masser y F. Salge (eds.): *European Geographic Information Infrastructures: Opportunities and Pitfalls* (pp. 153-167). Londres: Taylor and Francis.
- Onsrud, H., Campbell, J. y van Loenen, B. (2010). Towards Voluntary Interoperable Open access Licenses for the Global Earth Observation System of Systems (GEOSS). *International Journal of Spatial Data Infrastructure Research*, 5. Recuperado de https://works.bepress.com/harlan_onsrud/2.
- Open Knowledge Foundation (s. f.). En Wikipedia. Recuperado de https://es.wikipedia.org/wiki/Open_Knowledge_Foundation.
- Open Knowledge International (s. f.). Open Definition 2.0. *Open Definition*.

- Recuperado de <http://opendefinition.org/od/2.1/en>.
- Open Source Initiative (s. f.). *The Open Source definition*. Recuperado de <https://opensource.org/docs/osd>.
- Otegui Tellechea, J. (2012). *Quality and fitness-for-use assessments on ther primary data indexed at the Global Biodiversity Information Facility (GBIF)* (tesis doctoral). Recuperada de Depósito Académico Digital de la Universidad de Navarra (acceso: 28205). Recuperado de <http://dadun.unav.edu/handle/10171>.
- Oxenham, S. (2016). Legal confusion threatens to slow data science. *Nature*, 536(7614). Recuperado de <http://www.nature.com/news/legal-confusion-threatens-to-slow-data-science-1.20359>.
- Page, R. (2008). Biodiversity informatics: the challenge of linking data and the role of shared identifiers. *Briefing in Bioinformatics*, 9(5), 345-354. Recuperado de <http://bib.oxfordjournals.org/content/9/5/345.full>.
- Pavitt, K. (2001). Public Policies to Support Basic Research: What Can the Rest of the World Learn from US Theory and Practice? (And What They Should Not Learn). *Industrial and Corporate Change*, 10(3), 761-779. Recuperado de <https://academic.oup.com/icc/article-abstract/10/3/761/1276661/Public-Policies-to-Support-Basic-Research-What-Can?redirectedFrom=fulltext>.
- Pe'er, G., McNeely, J. A., Dieterich, M., Jonsson, B.G., Selva, N., Fitzgerald, J. M. y Nesshöver, C. (2013). IPBES: Opportunities and Challenges for SCB and Other Learned Societies. *Conservation Biology*, 27(1), 1-3. Recuperado de <http://onlinelibrary.wiley.com/doi/10.1111/cobi.12000/full>.
- Pereira, H. M., Navarro, L. M, y Martins, Í. S. (2012). Global Biodiversity Change: The Bad, the Good, and the Unknown. *Annual Review of Environment and Resources* 37 (1). 27. Recuperado de: <https://www.annualreviews.org/doi/pdf/10.1146/annurev-environ-042911-093511>
- Pereira, H. M., Ferrier, M., Walters, M. y Geller, N. (2013). Essential Biodiversity Variables. *Science*, 339(6117), 277-278. Recuperado de http://cba.fc.ul.pt/about/news/2013/Pereira_Science_PolicyForumDraft.pdf.
- Pereira, H. M., (2013). Essential Biodiversity Variables for Global Earth Observation. Firts Plenary of IPBES. Bonn (21-23 enero) Recuperado de: <https://www.slideserve.com/shaina/essential-biodiversity-variables-for-global-earth-observation>

- Peretti Pezzi, R. (2015). Open science: from hypertexts to hyperobjects. En S. Albagli, M. L. Maciel y A. H. Abdo (eds.): *Open science, open issues* (pp. 157-188). Río de Janeiro: IBICT; UNIRIO.
- Perkmann, M. y Walsh, K. (2009). The two faces of collaboration: impacts of university-industry relations on public research. *Industrial and Corporate Change*, 18(6), 1033-1065. Recuperado de [http://www.chss.uqam.ca/Portals/0/docs/sts8020/\(29\)Perkmann%20and%20Walsh.pdf](http://www.chss.uqam.ca/Portals/0/docs/sts8020/(29)Perkmann%20and%20Walsh.pdf).
- Perrings, C., Duraiappah, A., Larigauderie, A. y Mooney, H. (2011). The Biodiversity and Ecosystem Services Science-Policy Interface. *Science*, 331(6021), 1139-1140., Recuperado de <http://science.sciencemag.org/content/331/6021/1139.full>.
- Perspectiva Mundial sobre la Biodiversidad 3* (2010). Montreal: Secretaría del Convenio para la Biodiversidad Biológica.
- Peterson, A. T., Knapp, S., Guralnick, R., Soberón, J. y Holder, M. T. (2010). The big questions for biodiversity informatics. *Systematics and Biodiversity* 8(2): 159-168. Recuperado de <https://kuscholarworks.ku.edu/bitstream/handle/1808/13674/soberon2010QUESTIONS.pdf;sequence=1>.
- Peterson, A. T. y Soberón, J. y Krishtalka, L. (2015). A global perspective on decadal challenges and priorities in biodiversity informatics. *BMC Ecology*, 15(15), 2-9. Recuperado de <http://bmcecol.biomedcentral.com/articles/10.1186/s12898-015-0046-8>.
- Philips, A., Williamson, I. y Ezigbalike, C. (2012). Spatial data infrastructure concepts. *The Australian Surveyor*, 44(1), 20-28. Recuperado de <http://www.tandfonline.com/doi/abs/10.1080/00050326.1999.10441899>.
- Pigliucci, M. (2009). The end of theory in science? *Embo Report*, 10(6), 534. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2711825>.
- Pollock, N. y Williams, R. (2010). e-Infrastructures: How Do We Know and Understand Them? Strategic Ethnography and the Biography of Artefacts. *Computer Supported Cooperative Work*, 19(6), 521-556. Recuperado de <http://link.springer.com/article/10.1007/s10606-010-9129-4>.
- Proença, V., Martin, L. J., Pereira, H. M., Fernández, M., McRae, L., Belnap, J. *et al.* (2016). Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. *Biological Conservation*, 213, 256-263. Recuperado de

- <http://www.sciencedirect.com/science/article/pii/S0006320716302786>.
- Rands, M. R. W., Adams, W. A., Bennun, L., Butchart, S. H. M., Clements, A., Coomes, D. *et al.* (2010). Biodiversity Conservation: challenges beyond 2010. *Science*, 329(5997), 1298-1303 Recuperado de <http://science.sciencemag.org/content/329/5997/1298>.
- Raubitschek, J. y Latker, N. (2005). Reasonable pricing. A new twist for march-in rights under the Bayh Dole Act. *Santa Clara High Technology Law Journal*, 22(1), 149-167. Recuperado de <http://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1399&context=chtlj>.
- RDA-CODATA Legal Interoperability Interest Group (2016). *Legal Interoperability of Research Data: principles and implementation guidelines*. Recuperado de <https://zenodo.org/record/162241#.WHkI5VMrK71>.
- Redford, K., Groves, C., Medellín, R. A. y Robinson, J. G. (2012). Conservation Stories, Conservation Science, and the Role of the Intergovernmental Platform on Biodiversity and Ecosystem Services. *Conservation Biology*, 26(5), 757-759. Recuperado de <http://onlinelibrary.wiley.com/doi/10.1111/j.1523-1739.2012.01925.x/full>.
- Reichman, J. H. y Samuelson, P. (1997). Intellectual Property Rights in Data? *Vanderbilt Law Review*, 50(51), 51-166. Recuperado de <https://www.law.berkeley.edu/php-programs/faculty/facultyPubsPDF.php?facID=346&pubID=66>.
- Reichman, J. y Uhler, P. (1999). Database Protection at the Crossroads: Recent Development and Their Impact on Science and Technology. *Berkeley Technology Law Journal*, 14(2), 793-838. Recuperado de <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1249&context=btlj>.
- Research Data Alliance (s. f.). *About RDA*. Recuperado de <https://www.rd-alliance.org/about-rda>.
- (s. f.). *RDA/CODATA Legal Interoperability IG*. Recuperado de <https://www.rd-alliance.org/groups/rdacodata-legal-interoperability-ig.html>.
- (2015) “Semantic Interoperability”. Accessed March 2018. http://smw-rda.esc.rzg.mpg.de/index.php/Semantic_Interoperability. [Ref list]

- Research Information Network (2008). *To share or not to share: publication and quality assurance of research data outputs*. Recuperado de <http://www.rin.ac.uk/system/files/attachments/To-share-data-outputs-report.pdf>: RIN.
- y British Library (2009). *Patterns of information use and exchange: case studies and researchers in the Life Sciences*. Recuperado de http://www.rin.ac.uk/system/files/attachments/Patterns_information_use-REPORT_Nov09.pdf: Research Information Network & the British Library.
- Resnik, D. B. (2003). Strengthening the United States' Database Protection Laws. *Science and Engineering Ethics*, 9(3), 301-318. Recuperado de <https://link.springer.com/article/10.1007/s11948-003-0027-8>.
- Rizzo, G., Morando, F. y De Martín, J. C. (2009). Open data: la piattaforma di dati aperti per il Linked Data. *International Journal on Semantic Web and Information Systems*, 5(3), 1-22.
- ROARMAP (s. f.). Wellcome to ROARMAP. *Roarmap*. Recuperado de <http://roarmap.eprints.org>.
- Roberts, D. y Moritz, T. (2015). A framework for publishing primary biodiversity data. *BMC Bioinformatics*, 12(S15). Recuperado de <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-S15-I1>.
- Robertson, T. Döring, M., Guralnick, R., Bloom, D., Wieczorek, J., Braak, K., Otegui, *et al.* (2013). The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet. *PLoS ONE*, 9(8), e102623. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0102623>.
- Rodríguez, V. A. (2005). Tecnología y educación al resurgir del cuarto mundo. *HAOL*, 8, 175-182.
- Saarenmaa, H. (2005). *Sharing and Accessing Biodiversity Data*. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.6157&rep=rep1&type=pdf>.
- Sanbi (6 de abril de 2016). *Biodiversity Science*. Recuperado de <https://www.sanbi.org/biodiversity-science>.
- Santamaría, J. M. (3 de junio de 2017). La importancia de la ciencia ciudadana. *El Español*, Recuperado de http://www.elespanol.com/opinion/tribunas/20170602/220847915_12.htm

- Sarkar, I. N. (2007). Biodiversity informatics: organizing and linking information across the spectrum of life. *Briefings in Bioinformatics*, 8(5), 347-357. Recuperado de <http://bib.oxfordjournals.org/content/8/5/347.abstract>.
- Savage, C. J. y Vickers, A. J. (2009). Empirical Study of Data Sharing by Authors Publishing in PLoS Journal. *PLoS ONE*, 4(9), e7078. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007078>.
- Sayogo, D. S. y Pardo, T. (2011). Understanding the Capabilities and Critical Success. The case of DataONE. En *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 74-83). Nueva York: AMC. Recuperado de <http://dl.acm.org/citation.cfm?id=2037568>.
- Schaeffer B., Baranski B., Foerster T., and Brauner J. 2012. "A Service-oriented Framework for Real-time and Distributed Geoprocessing." In *Geospatial Free and Open Source Software in the 21st Century*, edited by E. Bocher and M. Neteler, 3–20. Heidelberg: Springer. doi:10.1007/978-3-642-10595-1_1.
- Schalk, P. (1998). Management of marine natural resources through by biodiversity informatics. *Marine Policy*, 22(3), 269-280. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0308597X9800013X>.
- Scharmer, O. (2016). *Theory U: Leading from the Future as it Emerges* (2.^a ed.). San Francisco: BK Currents book.
- Schimel, D., Hargrove, W., Hoffman, F. y MacMahon, J. (2007). NEON: a hierarchically designed national ecological network. *Frontiers in Ecology and the Environment*, 5(2), 59. Recuperado de [http://onlinelibrary.wiley.com/doi/10.1890/1540-9295\(2007\)5\[59:NAHDNE\]2.0.CO;2/full](http://onlinelibrary.wiley.com/doi/10.1890/1540-9295(2007)5[59:NAHDNE]2.0.CO;2/full).
- Schmeller, D. S., Mihoub, J. B., Arvanitidis, C., Costello, M. J., Fernández, M., Geller, G. N. *et al.* (2017). An operational definition of essential biodiversity variables. *Biodiversity and Conservation*, 26(12), 2967-2972. Recuperado de <https://link.springer.com/article/10.1007%2Fs10531-017-1386-9>.
- Schmidt, B. y Kuchma, I. (2012). Implementing Open access mandates in Europe. Gotinga: Universitätsverlag Göttingen. Recuperado de <https://www.univerlag.uni-goettingen.de/handle/3/isbn-978-3-86395-095-8>.
- Schmidt, B., Gemeinholzer, B. y Treloar, A. (2016). Open data in global Environmental Research: The Belmont Forum's Open data Survey. *PLoS ONE*, 11(1), e0146695.

- Schneider, M. (1998). The European Union Database Directive. *Berkeley Technological Law Journal*, 13(1), 551-564. Recuperado de <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1189&context=btlj>.
- Scholes, B. (2012). *What is an Essential Biodiversity Variable?* Frascati: GEOBON. Recuperado de https://earthobservations.org/documents/meetings/201202_geobon_ebv/07_ebv_what_is_an_ebv_bob_scholes.pdf.
- Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jürgens, N., Larigauderie, A. *et al.* (2008). Toward a Global Biodiversity Observing System. *Science* 321, 1044-1045. Recuperado de https://earthobservations.org/documents/cop/bi_geobon/200808_science_toward_a_global_biodiversity_observing_system.pdf.
- Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H. R., Tuama, O. E., Faith, D., P. *et al.* (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, 4(1), 139-146. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1877343511001394>.
- Science Common (s. f.). *Protocol for Implementing Open Access Data*. Recuperado de <http://sciencecommons.org/projects/publishing/open-access-data-protocol>.
- Secretaría del Convenio de Diversidad Biológica (2005). *Handbook on the Convention of Biological Diversity*. Montreal: Taylor & Francis.
- (2010). *Perspectiva Mundial sobre la Diversidad Biológica 3*. Montreal. Recuperado de <https://www.cbd.int/doc/publications/gbo/gbo3-final-es.pdf>.
- Secretariat of the Convention on Biological Diversity (2005). *Handbook of the Convention on Biological Diversity Including its Cartagena Protocol on Biosafety* (3.^a ed.). Montreal: s. e. Recuperado de <https://www.cbd.int/doc/handbook/cbd-hb-all-en.pdf>.
- Seidel, E. (10 de mayo de 2010). Scientists seeking NSF founding will soon be required to submit data management plans. *National Science Foundation Press Release*. Recuperado de http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928.
- Shaon, A. y Woolf, A. (2011). Long-term preservation for Spatial Data Infrastructures: a Metadata Framework and Geo-portal implementation. *D-Lib Magazine*, 17(9-10). Recuperado de <http://www.dlib.org/dlib/september11/shaon/09shan.html>.

- Shoiti Ikematu, R. (2001). Gestao de Metadados: Sua Evolução na Tecnologia da Informação. *Data Grama Zero. Revista de Ciência da Informação*, 2(6), A02-0. Recuperado de http://www.datagramazero.org.br/dez01/Art_02.htm.
- Silva, M. (2004). Bioinformatics, the clearing-house mechanism and the Convention on Biological Diversity. *Biodiversity Informatics*, 1, 23-29. Recuperado de <https://journals.ku.edu/index.php/jbi/article/view/6/4>.
- Singh Khalsa, S. J., Nativi, S. y Geller, G. N. (2009). The GEOSS Interoperability Process Pilot Project (IP3). *IEEE Transactions on Geoscience and Remote Sensing* 47(1). Recuperado de <http://ieeexplore.ieee.org/document/4694112/?arnumber=4694112&tag=1>
- Soberón, J. y Peterson, A. T. (2004). Biodiversity informatics: managing and applying primary biodiversity data. *The Royal Society*, 359, 689-698. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1693343/pdf/15253354.pdf>.
- Soberón, J. y Sarukhán, J. R. (2010). A new mechanism for science-policy transfer and biodiversity governance? *Environmental Conservation* 36(4), 265-267. Recuperado de <https://www.cambridge.org/core/services/aop-cambridge-core/content/view/S0376892910000226>.
- Soberón, J. y Townsend, P. (2015). Biodiversity Governance: A Tower of Babel of Scales and Cultures. *PLoS Biology*, 13(3), e1002108. Recuperado de <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002108>.
- Spek, R. V. y Spijkervet, A. (1997). *Knowledge Management: Dealing Intelligently with Knowledge*. Utrecht: Knowledge Management Network.
- Spierenburg, M. (2012). Getting the Message Across: Biodiversity Science and Policy Interfaces. A Review. *GALA* 21(2), 125-134. Recuperado de <http://www.spiral-project.eu/sites/default/files/Spierenburg.pdf>.
- Star, S. L. y Ruhleder, K. (1996). Steps toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, 7(1), 111-134. Recuperado de <http://dx.doi.org/10.1287/isre.7.1.111>.
- Stenseke, M. y Larigauderie, A., (2017). The role, importance and challenges of social sciences and humanities in the work of the intergovernmental science-policy platform on biodiversity and ecosystem services (IPBES). *Journal Innovation: The European Journal of Social Science Research Volume 31, 2018 - Issue sup1: Supplement: The Social Sciences and Humanities in the Intergovernmental Science-*

- Policy Platform on Biodiversity and Ecosystem Services (IPBES)*. Recuperado de <https://www.tandfonline.com/doi/abs/10.1080/13511610.2017.1398076>.
- Stodden, V. (2009). The Legal Framework for Reproducible Research in the Sciences: Licensing and Copyright. *Computing in Science & Engineering*, 11, 35-40. Recuperado de <http://web.stanford.edu/~vcs/papers/ReproducibleResearchLegal08172008.pdf>.
- Stone, R. (2010). Earth-Observation Summit Endorses Global Data Sharing. *Science*, 330(6006), 902. Recuperado de <http://science.sciencemag.org/content/330/6006/902.full>.
- Suber, P. (2012). *Open access*. Cambridge (Massachussets); Londres: The MIT Press. Recuperado de https://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version.pdf.
- Sudmanns, M., Tiede, D., Lang, S., & Baraldi, A. (2018). Semantic and syntactic interoperability in online processing of big Earth observation data. *International Journal of Digital Earth*, 11(1), 95–112. <http://doi.org/10.1080/17538947.2017.1332112>.
- Sweat of the brow (s f.). En Wikipedia. Recuperado de http://en.wikipedia.org/wiki/Sweat_of_the_brow.
- Takacs, D. (1996). *The Idea of Biodiversity: Philosophies of Paradise*. Baltimore: The Johns Hopkins University Press.
- Tenopir C., Dalton, E. D., Allard, S., Frame, M. Pjesivac, I., Birch, B. *et al.* (2015). Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLoS One*, 10(8), e0134826. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134826>.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M. y Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE*, 6(6), e21101. Recuperado de <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021101>.
- The Digital Archiving Consultancy Limited (2008). *Final Report for DG Information Society and Media, Unit F - Géant and e-Infrastructure.Towards a European e-Infrastructure for e-Science Digital Repositories*. Middlesex: The Digital Archiving Consultancy Limited.

- Knappenberger, B. (productor y director) (2014). *The Internet's own boy. The Story of Aaron Swartz* [película]. Estados Unidos: FilmBuff y Participant Media
- The Royal Society Intellectual Property Group (2003). Keeping Science Open: The Effects of Intellectual Property Policy on the Conduct of Science. *Royal Society* Recuperado de https://royalsociety.org/~media/Royal_Society_Content/policy/publications/2003/9845.pdf:.:
- Tinker, C. (1995). A «new breed» of treaty: the United Nations Convention on Biological Diversity. *PACE Environmental Law Review*, 13(1), 191-218 Recuperado de <http://digitalcommons.pace.edu/cgi/viewcontent.cgi?article=1397&context=pehr>.
- Tolba, M. K. (1998). *Global Environmental Diplomacy: Negotiating Environmental Agreements for the World, 1973-1992*. Cambridge (Massachusetts): The MIT Press.
- Toledo, V. (1994). La diversidad biológica de México: nuevos retos para la investigación en los noventas. *Ciencias*, 34, 43-59. Recuperado de <http://www.ejournal.unam.mx/cns/no34/CNS03407.pdf>.
- Torres-Salinas, D., Robinson-García, N. y Cabezas-Clavijo, A. (2012). Compartir los datos de investigación en ciencia: introducción al Data Sharing. *El profesional de la información*, 21(2), 173-184. Recuperado de <http://eprints.rclis.org/16786/1/data%20sharing.pdf>.
- Tréguer, F. (2012). Interoperability Case Study The European Union as an Institutional Design for Legal Interoperability. *The Berkman Center for Internet and Society at Harvard University*, 18(2012). Recuperado de http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2148543.
- Tuomi, I. (1999). Data is More than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory. En *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, HICSS-32*. Washington, DC: IEEE Computer Society. Recuperado de <http://ieeexplore.ieee.org/abstract/document/772795>.
- Tucker, M. A. *et al.* 2018. Moving in the Anthropocene: Global reductions in terrestrial mammalian movements *Science* 26 Jan 2018. Vol. 359, Issue 6374, pp. 466-469. DOI: 10.1126/science.aam9712
- Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z. *et al.* (2015). Free and open-access satellite data are key to biodiversity

- conservation. *Biological Conservation*, 182, 173-176. Recuperado de <https://doi.org/10.1016/j.biocon.2014.11.048>).
- Turnhout, E., Neves, K. y Lijster, E. de (2014). 'Measurementality' in biodiversity governance: knowledge, transparency, and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). *Environment and Planning A: Economy and Space*, 46(3), 581-597. Recuperado de <http://journals.sagepub.com/doi/pdf/10.1068/a4629>.
- Uhlir, P. (2015). The Value of Open data Sharing [informe preparado para el GEO-XII Plenary por GEO Participating Organization CODATA, ICSU Committee on Data for Science and Technology]. Recuperado de https://www.earthobservations.org/documents/dsp/20151130_the_value_of_open_data_sharing.pdf.
- (2016). *Why Legal Interoperability? Purpose and Key Concepts*. Conference Paper. Parde de la sesión realizada por Robert Chen en SciDATACOM 2016. Recuperado de <http://www.scidatacon.org/2016/sessions/84/>.
- y Reichman, J. H. (2003). A contractually reconstructed research commons for scientific data in a highly protectionist intellectual property environment. *Law and Contemporary Problems*, 66(2003), 315-462. Recuperado de http://scholarship.law.duke.edu/faculty_scholarship/894.
- y Schröder, P. (2007). Open data for global science. *Data Science Journal*, 6, 36-53. Recuperado de <http://datascience.codata.org/articles/abstract/10.2481/dsj.6.OD36>.
- Chen, R. S., Gabrynowicz, J. I. y Janssen, K. (2009). Toward Implementation of the Global Earth Observation System of System Data Sharing principles. *Journal of Space Law*, 35(1), 201-290.
- Chen, R. S. y Alonso García, E. (2016). *Issues in the Legal Interoperability of Research Data*. Conference Paper. Recuperado de <http://www.scidatacon.org/2016/sessions/84/>).
- Usher, M. B. (1991). Biodiversity a scientific challenge for resource managers in the 1990s. En F. D. Pineda *et al.* (eds.): *Diversidad biológica* (pp. 33-40). Madrid: Fundación Ramón Areces.
- Vadrot, A. (2014). *The politics of knowledge and global biodiversity*. Abingdon: Routledge.
- Valoir, T. (2000). Government funded inventions, the Bayh-Dole Act and the Hopkins v. Cell Pro march-in rights controversy. *Texas intellectual Property Law Journal*, 8(2), 211-239.

- Van den Hove, S. (2007). A rationale for science-policy interfaces. *Futures*, 39(7), 807-826.
- y Chabason, L. (2009). The Debate on an intergovernmental science-policy platform on biodiversity and ecosystem services (IPBES): exploring gaps and needs. Discussion Paper 1/2009. *Institut du développement durable et des relations internationales*. Recuperado de http://www.iddri.org/Publications/Collections/Idees-pour-le-debat/Id_090104_gap_analysis-4Feb.pdf.
- Van der Veer, H. Wiles, A. (2008). ETSI White Paper No. 3 Achieving Technical Interoperability - the ETSI Approach. Recuperado de: <https://portal.etsi.org/CTI/Downloads/ETSIApproach/IOP%20whitepaper%20Edition%203%20final.pdf>
- Vercelli, A. (2004). *La conquista silenciosa del ciberespacio: Creative Commons y el diseño de entornos digitales como nuevo arte regulativo en Internet*. Buenos Aires: Ariel Hernán Vercelli. Recuperado de <http://eprints.rclis.org/9926/1/lcsdcvercelli.pdf>.
- Vergara, G. P. (13 de octubre de 2011). Las claves sobre IPBES, el IPCC de la biodiversidad [entrada de blog]. *Asuntos Verdes*. Recuperado de <http://asuntosverdes.blogspot.com.es/2011/10/las-claves-sobre-ipbes-el-ipcc-de-la.html>.
- Vickers, A. J. (2006). Whose data set is it anyway? Sharing raw data from randomized trials. *Trials*, 7, 15. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1489946>.
- Vogel, J. H. (2005). Sovereignty as a Trojan Horse: How the Convention on Biological Diversity morphs Biopiracy into Biofraud. En B. Hocking (ed.): *Unfinished constitutional business? Rethinking indigenous self-determination* (p. 230). Camberra: Aboriginal Studies Press.
- Vohland, K., Mlambo, M. C., Horta, L. D., Jonsson, B., Paulsch, A., Martínez, S. I. (2011). How to ensure a credible and efficient IPBES? *Environnement Science Policy*, 14(8), 1188-1194. Recuperado de <http://www.sciencedirect.com/science/article/pii/S1462901111001389>.
- Vollmar, A., Macklin, J. A. y Ford, L. S. (2010). Natural history specimen digitation: challenges and concerns. *Biodiversity informatics*, 7(2), 93-112. Recuperado de <https://journals.ku.edu/index.php/jbi/article/view/3992/3806>.
- Vollmer, T. (6 de septiembre 2016). Commission proposes to limit text and data mining in Europe [entrada de blog]. *Communika*. Recuperado de

- <http://www.communia-association.org/2016/09/06/commission-proposes-limit-text-data-mining-europe>.
- Von Hippel, E. y Krogh, V. (2003). Open source *software* and the «private-collective» innovation model: Issues for organization science. *Organization Science*, 14(2), 209-223. Recuperado de <https://pubsonline.informs.org/doi/abs/10.1287/orsc.14.2.209.14992>.
- Walpole, M. *et al.* (2009). Trackin progress toward the 2010 biodiversity target and beyond. *Science*, 325, 1503-1504. Recuperado de <http://science.sciencemag.org/content/325/5947/1503>.
- Watson, R. T. (2001). Turning science into policy: challenges and experiences from the science-policy interface. *Philosophical transactions of the Royal Society of London. Series B. Biological sciences*, 360(1454), 471-477. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1569452>.
- Weber, R. H. (2014). *Legal Interoperability as a Tool for Combatting Fragmentation*. Ontario; Londres: Centre for International Governance Innovation and the Royal Institute of International Affairs (Paper Series, 4). Recuperado de https://www.cigionline.org/sites/default/files/gcig_paper_no4.pdf.
- Wetzel, F. T., Saarenmaa, H., Regan, E., Martin, C. S., Mergen, P., Smirnova, L. *et al.* (2015). The roles and contributions of Biodiversity Observation Networks (BONs) in better tracking progress to 2020 biodiversity targets: a European case study. *Biodiversity* 16(2 y 3), 137-149. Recuperado de <http://www.tandfonline.com/doi/full/10.1080/14888386.2015.1075902>.
- Wheeler, Q. D. (2004). What is GBIF? *BioScience*, 54(8), 717. Recuperado de <http://bioscience.oxfordjournals.org/content/54/8/717.short>.
- Raven, P. H. y Wilson, E. O. (2004). Taxonomy: Impediment or Expedient? *Science*, 303(5656), 285. Recuperado de http://science.sciencemag.org/content/303/5656/285?ijkey=1443125e92b286e39bc7b6bc1ad987b6e1bb7438&keytype2=tf_ipsecsha.
- Wilbanks, J. (2006). Another reason for opening access to research. *British Medical Association*, 333(2006), 1306-1308. Recuperado de <http://pubmedcentralcanada.ca/pmcc/articles/PMC1761190/pdf/bmj-333-7582-inspiration-01306.pdf>.
- (2008). Public domain, copyright licenses and the freedom to integrate science. *Journal of Science Communication*, 7(2), 1-10. Recuperado de [https://jcom.sissa.it/archive/07/02/Jcom0702\(2008\)C01/Jcom0702\(2008\)C04](https://jcom.sissa.it/archive/07/02/Jcom0702(2008)C01/Jcom0702(2008)C04).

- (2011). Openness as infrastructure. *Journal of Cheminformatics*, 3(36). Recuperado de <http://jcheminf.springeropen.com/articles/10.1186/1758-2946-3-36>.
- Wilkinson, M. D., Verborgh, R., Da Silva Santos, L. O. B., Clark, T., Swertz, M. A., Kelpin, F. D. L. *et al.* (2016). Interoperability and FAIRness through a novel combination of Web technologies. *PeerJ Computer Science*, 3, e110. Recuperado de <https://peerj.com/articles/cs-110>.
- Wilkinson, M. D., Dumortier, M. y Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(2016), artículo 160018. Recuperado de <http://dx.doi.org/10.1038/sdata.2016.18>.
- Willinsky, J. (1999). *Technologies of Knowing: A Proposal for the Human Sciences*. Boston: Beacon Press.
- (2006). *The Access Principle: The Case for Open access to Research and Scholarship*. Cambridge (Massachusetts): The MIT Press.
- Wilson, E. O. (ed.) (1988). *Biodiversity*. Washington: National Academic Press.
- Wojick, D. (6 diciembre de 2016). The transition team [entrada de blog]. *Poynder*. Recuperado de <http://poynder.blogspot.pt/2016/12/tracking-trump.html>.
- Xalabarder Plantada, R. (2006). Les llicències Creative Commons: una alternativa al copyright? *UOC Papers: revista sobre la societat del coneixement*, 2(2006), 4-12. Recuperado de http://openaccess.uoc.edu/webapps/o2/bitstream/10609/7442/1/Xalabarder_llicenciescommons_2006_cat.pdf.
- Yang, X. y Wang, L. (2011). *Guide to eScience: Next generation Scientific Research and Discovery*. Londres: Springer Science & Business Media.
- Yesson, C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M. *et al.* (2007). How Global Is the Global Biodiversity Information Facility? *PLoS ONE*, 2(11), e1124. Recuperado de <http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0001124&type=printable>.

NORMAS Y RECOMENDACIONES

Estados Unidos

Normas

- S. 1983 — 93rd Congress: Endangered Species Conservation Act.
www.GovTrack.us. 1973. April 9, 2018
<https://www.govtrack.us/congress/bills/93/s1983>
- S. 2037 — 105th Congress: Digital Millennium Copyright Act of 1998.
www.GovTrack.us. 1998. April 7, 2018
<https://www.govtrack.us/congress/bills/105/s2037>
- S. 2079 — 96th Congress: A bill to improve the administration of the patent and trademark laws by establishing the [...] www.GovTrack.us. 1979. April 7, 2018
<https://www.govtrack.us/congress/bills/96/s2079>
- H.R. 2652 — 105th Congress: Collections of Information Antipiracy Act.
www.GovTrack.us. 1997. April 7, 2018
<https://www.govtrack.us/congress/bills/105/hr2652>
- H.R. 1858 — 106th Congress: Consumer and Investor Access to Information Act of 1999. www.GovTrack.us. 1999. April 7, 2018
<https://www.govtrack.us/congress/bills/106/hr1858>
- H.R. 4577 — 106th Congress: Consolidated Appropriations Act, 2001.”
www.GovTrack.us. 2000. April 7, 2018
<https://www.govtrack.us/congress/bills/106/hr4577>
- H.R. 2272 — 110th Congress: America COMPETES Act. www.GovTrack.us. 2007.
April 7, 2018 <https://www.govtrack.us/congress/bills/110/hr2272>
- H.R. 5253 — 111th Congress: Federal Research Public Access Act of 2010.
www.GovTrack.us. 2010. April 9, 2018
<https://www.govtrack.us/congress/bills/111/hr5253>
- S. 779 — 114th Congress: Fair Access to Science and Technology Research Act of 2015.
www.GovTrack.us. 2015. April 7, 2018
<https://www.govtrack.us/congress/bills/114/s779>

Jurisprudencia

- Feist Publications v. Rural Telephone Service Co., 499 U.S. 340, 111 S. Ct. 1282, 113 L. Ed. 2d 358, 1991 U.S. Recuperado de <http://caselaw.findlaw.com/us-supreme-court/499/340.html>.

Informes

White House (1971). Memorandum and Statement of Government Patent Policy, Washington. Recuperado de <http://www.presidency.ucsb.edu/ws/?pid=3130>.

National Science Foundation (1995). On the Full and Open Exchange of Scientific Data. Recuperado de <https://www.nap.edu/catalog/18769/on-the-full-and-open-exchange-of-scientific-data>.

US Copyright Office (1997). Report on legal protection for databases, s.l.: U.S. Copyright Office. Recuperado de <https://www.copyright.gov/reports/db4.pdf>.

National Institute of Health (2003). NIH Data Sharing Policy and Implementation Guidance. Recuperado de https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm.

National Science Foundation (2007). Ciberinfrastructures vision for 21st Century Discovery. Recuperado de https://www.nsf.gov/pubs/2007/nsf0728/nsf0728_2.pdf.

National Science Foundation, Office of Cyberinfrastructure, Directorate for Computer & Information Science & Engineering (2008). Sustainable Digital Data Preservation and Access Network Partners (DataNET). Recuperado de <https://www.nsf.gov/pubs/2007/nsf07601/nsf07601.htm>.

National Science Foundation (2010). Data Management for NSF SBE Directorate. Recuperado de: https://www.nsf.gov/sbe/SBE_DataMgmtPlanPolicy.pdf.

Executive Office of the President - Office of Science and Technology Policy (2013). Memorandum for the heads of Executive Departments and agencies. Recuperado de https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

National Science Foundation (2013). Chapter VI - Other Post Award Requirements and Considerations. Recuperado de https://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/aag_6.jsp#VID4.

National Science Foundation (2016). www.nsf.gov/about/. Recuperado de <http://www.nsf.gov/about/>.

National Science Foundation, s.f. Data Archiving Policy. Recuperado de <https://www.nsf.gov/sbe/ses/common/archive.jsp>.

Europa

Normas

Unión Europea. Directiva (UE) 1995/46 del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. Diario Oficial de la Unión Europea L 281, 23 de noviembre de 1995. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:31995L0046>.

Unión Europea. Directiva (UE) 1996/9 del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos, Diario Oficial de la Unión Europea L 77, 27 de marzo de 1996. Recuperado de <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=LEGISSUM%3A126028>.

Unión Europea. Directiva (UE) 2001/29 del Consejo, de 22 de mayo de 2001, relativa a la armonización de determinados aspectos de los derechos de autor y derechos afines a los derechos de autor en la sociedad de la información. Diario Oficial de la Unión Europea L 167/10, 22 de junio de 2001. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32001L0029&from=ES>.

Unión Europea. Directiva (UE) 2003/4 del Consejo, de 28 de enero, relativa al acceso público a la información medioambiental y por la que se deroga la Directiva 90/213/CE, Diario Oficial de la Unión Europea L 041, 14 de febrero de 2003 p. 0026 – 0032. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:32003L0004>.

Unión Europea. Directiva (UE) 2003/98 del Consejo, de 17 de noviembre, sobre la reutilización de la información del sector público (Directiva PSI). Diario Oficial de la Unión Europea L 345, 31 de diciembre de 2003, p. 0090 – 0096. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32003L0098&from=ES>.

Unión Europea. Directiva (UE) 2007/2 del Consejo, de 14 de marzo, por la cual se establece una infraestructura de información espacial en la Comunidad Europea, (INSPIRE). Diario Oficial de la Unión Europea L 212, 25 de abril de 2007. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32007L0002>.

España. Ley 42/2007, de 13 de diciembre, del Patrimonio Natural y de la Biodiversidad. Boletín Oficial del Estado, 14 de diciembre de 2007, núm. 299. Recuperado de <https://www.boe.es/buscar/act.php?id=BOE-A-2007-21490>.

Reglamento (UE) 723/2009, de 23 de junio de 2009, relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC). Diario Oficial de la Unión Europea L 206/1, 8 de agosto de 2009. Recuperado de <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:206:0001:0008:ES:PDF>.

Unión Europea. Comunicación (UE) 2011/0882 de la Comisión al Parlamento Europeo, al Consejo, al Comité económico y social europeo y al Comité de las Regiones (2011) Datos abiertos, un motor para la innovación, el crecimiento y la gobernanza transparente. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52011DC0882>.

España. Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación. Boletín Oficial del Estado, 2 de junio de 2011, núm. 131, Recuperado de <https://www.boe.es/buscar/act.php?id=BOE-A-2011-9617&p=20171007&tn=0>.

Unión Europea. Directiva (UE) 2013/37, de 26 de junio que entró en vigor el 17 de julio de 2013 por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. Recuperado de <https://www.boe.es/doue/2013/175/L00001-00008.pdf>.

Reglamento (UE) 1261/2013, de 2 de diciembre de 2013, por el que se modifica el Reglamento (UE) n°. 723/1999 relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC). Diario Oficial de la Unión Europea L 326/1, 6 de diciembre de 2013. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32013R1261&from=EN>.

Unión Europea. Comunicación (UE) 2014/442 de la Comisión al Parlamento europeo, al Consejo, al Comité económico y social europeo y al Comité de las Regiones. Hacia una economía de datos próspera. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=COM:2014:442:FIN&rid=1>.

Reglamento (UE) 2016/679, de 27 de abril, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos L 119/1, 4 de mayo de 2016. Recuperado de:

<https://publications.europa.eu/es/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-es>

Unión Europea. Directiva (UE) 2016/943 del Consejo, de 8 de junio, relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas. Diario Oficial de la Unión Europea L 157/1, 15 de junio de 2016. Recuperado de <https://publications.europa.eu/en/publication-detail/-/publication/dc0eda0f-32bc-11e6-b497-01aa75ed71a1/language-es>.

Unión Europea. Decisión de ejecución (UE) 2017/499 de la Comisión, de 17 de marzo de 2017, relativa a la creación de la Infraestructura Virtual Europea de Ciencia y Tecnología para la Investigación de la Biodiversidad y de los Ecosistemas — Consorcio de Infraestructuras de Investigación Europeas (LifeWatch ERIC). Diario Oficial de la Unión Europea L 76/35, 22 de marzo de 2017. Recuperado de http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=uriserv:OJ.L_.2017.076.01.0035.01.SPA&toc=OJ:L:2017:076:TOC.

Informes

Comisión Europea (1988). Green Paper on Copyright and the Challenge of Technology. Recuperado de <http://aei.pitt.edu/1209/>

Comisión Europea (2007). Annual Information Society Report. Recuperado de http://aei.pitt.edu/45834/1/com2007_0146.pdf.

Comunidades Europeas (2008). European Roadmap for Research Infrastructures. Roadmap. Recuperado de http://www.idi.mineco.gob.es/stfls/MICINN/Investigacion/FICHEROS/Internacional/esfri_roadmap_update_2008.pdf.

Comisión Europea (2010). Digital Agenda for Europe. Recuperado de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=LEGISSUM:si0016>.

Comisión Europea (2012). European Commission background note on open access to publications and data in Horizon 2020. Recuperado de http://ec.europa.eu/research/science-society/document_library/pdf_06/background-paper-open-access-october-2012_en.pdf.

Comisión Europea (2014). The European Union explained: Digital Agenda for Europe. Recuperado de http://eige.europa.eu/resources/digital_agenda_en.pdf.

Committee on Legal Affairs (2015). Draft Report: on the implementation of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (2014/2256(INI)). Recuperado de https://pub.juliareda.eu/copyright_evaluation_report.pdf.

Comisión Europea (2016). Guidelines on Open access to Scientific Publications and Research Data in Horizon 2020. Recuperado de http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

Comisión Europea (2017). Open access in Horizon 2020. Recuperado de https://ec.europa.eu/research/press/2016/pdf/opendata-infographic_072016.pdf.

Comisión Europea, s.f. Open access Pilot in FP7. Recuperado de http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-pilot_en.pdf

Internacional

Acuerdos/tratados

Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio (1994). Marrakech: World Trade Organization. Recuperado de https://www.wto.org/SPANISH/DOCS_S/legal_s/27-trips.pdf.

Convenio de Berna para la Protección de las Obras Literarias y Artísticas (1886). Ginebra: Organización Mundial de la Propiedad Intelectual. Recuperado de http://www.wipo.int/treaties/es/text.jsp?file_id=283698.

Tratado Antártico (1959). Washington DC: Secretaría del Tratado Antártico.

Convención relativa a los Humedales de importancia Internacional Especialmente como Hábitat de Aves Acuáticas (1971). París: Convención relativa a los Humedales de importancia Internacional Especialmente como Hábitat de Aves Acuáticas. Recuperado de https://www.ramsar.org/sites/default/files/documents/library/current_convention_s.pdf.

Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres (CITES) (1973). Washington: Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres. Recuperado de <https://cites.org/esp/disc/text.php>.

- Convención sobre la diversidad biológica*, (1992). Washington: Convenio de diversidad biológica. Recuperado de http://www.acnur.org/t3/fileadmin/Documentos/Pueblos_indigenas/convencion_diversidad_biologica_1992.pdf?view=1.
- Cumbre de la Tierra. Río de Janeiro* (1992). Río de Janeiro: Cumbre de la Tierra. Recuperado de http://www.unesco.org/education/pdf/RIO_S.PDF.
- Convención Marco de Naciones Unidas sobre Cambio Climático*, (1994). Washington: Convención Marco de Naciones Unidas sobre Cambio climático. Recuperado de <http://unfccc.int/resource/docs/convkp/convsp.pdf>.
- Protocolo de Kioto* (2005). Kioto Protocolo de Kioto. Recuperado de <https://unfccc.int/resource/docs/convkp/kpspan.pdf>.
- Convención de las Naciones Unidas para la lucha contra la Desertificación* (1994). París. Convención de las Naciones Unidas para la lucha contra la Desertificación. Recuperado de https://www2.unccd.int/sites/default/files/relevant-links/2017-08/UNCCD_Convention_text_SPA.pdf.

Informes

- Committee on Issues in the Transborder Flow of Scientific Data of the National Research Council (1997). Bits of Power: Issues in Global Access to Scientific. Recuperado de <https://www.nap.edu/catalog/5504/bits-of-power-issues-in-global-access-to-scientific-data>.
- Diversitas (2005). Oaxaca Declaration on Biodiversity. Recuperado de www.diversitasinternational.org/activities/open-science-conferences/osc1-1/oaxacadeclaration.
- FAO (2009). International Treaty on Plant Genetic Resources for Food and Agriculture. Recuperado de <http://www.fao.org/3/a-i0510e.pdf>.
- Foro de Alto Nivel (2005). Declaración de París sobre la eficacia de la ayuda al desarrollo.
- GBIF (2010). GBIF Memorandum of Understanding. Recuperado de <http://www.gbif.org/resource/80661>.
- GBIF (2012). GBIF and IPBES: leveraging a decade of investment in biodiversity data sharing. Recuperado de: <http://prodapi-vh.gbif.org/resource/80740>.
- GBIF (2015). Global Biodiversity Informatics Outlook: Delivering Biodiversity Knowledge in the Information Age. Recuperado de:

<https://www.gbif.org/document/80859/global-biodiversity-informatics-outlook-delivering-biodiversity-knowledge-in-the-information-age>

GEO Secretariat (2005). Global Earth Observations System of Systems 10-Year Implementation Plan Reference Document. Recuperado de <https://www.earthobservations.org/documents/10-Year%20Plan%20Reference%20Document.pdf>.

GEO (2008). The GEO Biodiversity Observation Network Concept Document, Recuperado de http://www.geobon.org/Downloads/Other_documents/200811_geobon_concept_document.pdf.

GEO-V (2008). GEO BON Implementation Overview: Early products as a vision for building the network, Recuperado de http://www.earthobservations.org/documents/geo_v/07_%20GEO%20Bon%20-%20Implementation%20Overview.pdf

GEO (2010). GEOSS: Data Sharing Accessing Plan, Reuperado de https://www.earthobservations.org/documents/geo_vii/07_GEOSS%20Data%20Sharing%20Action%20Plan%20Rev2.pdf.

GEO (2010). Group on Earth Observations Biodiversity Observation Network (GEO BON). Principles of the GEO BON Information Architecture. Recuperado de: https://www.earthobservations.org/documents/cop/bi_geobon/geobon_information_architecture_principles.pdf.

GEO (2012). GEO 2012-2015 work plan. Recuperado de http://www.earthobservations.org/documents/work%20plan/GEO%202012-2015%20Work%20Plan_Rev2.pdf.

GEO (2013). A Working Meeting for the vision and architecture of GEOSS Information System, Frascati: https://www.earthobservations.org/documents/se/130_in_board_vision_and_architecture_of_geoss_information_system.pdf.

GEO Data Sharing Working Group (DSWG) of the Group on Earth Observations (2014). GEOSS Data Sharing Principles post 2015. Recuperado de http://www.earthobservations.org/documents/dswg/10_GEOSS%20Data%20Sharing%20Principles%20post%202015.pdf.

GEO Data Sharing Working Group (2014). White Paper: Mechanisms to Share Data as Part of GEOSS Data-CORE. Recuperado de: <https://www.earthobservations.org/documents/dswg/Annex%20VI%20->

%20Mechanisms%20to%20share%20data%20as%20part%20of%20GEOSS%20Data_CORE.pdf.

GEO-X (2014). Sesión plenaria celebrada en Ginebra. Group on Earth Observation. Recuperado de <http://www.earthobservations.org/geo10.php>.

GEO-XI (2014). Towards GEO Data Management Principles (Report of the Data Management Principles Task Force). Recuperado de: https://www.earthobservations.org/documents/geo_xi/GEO-XI_07_Towards%20GEO%20Data%20Management%20Principles-Report%20of%20the%20Data%20Management%20Principals%20Task%20Force.pdf.

GEO XIII (2014). The value of Open data, Recuperado de http://www.earthobservations.org/documents/dsp/20151130_the_value_of_open_data_sharing.pdf.

GEO Data Management Principles Task Force (2015). GEOSS Data Management Principles. Recuperado de: https://www.earthobservations.org/documents/dswg/201504_data_management_principles_long_final.pdf

GEO BON Management Committee (2017). GEO BON Strategy for development of Essential Biodiversity Variables. Recuperado de: www.geobon.org/downloads/Other_documents/Essential_Biodiversity_Variable_Strategy_v2.pdf

GEO BON, s.f. How GEO BON can contribute to the conservation and the sustainable use of biological diversity. Recuperado de http://www.earthobservations.org/documents/cop/bi_geobon/201004_geobon_brochure.pdf.

G8 Science Minister's (2013). G8 Science Ministers' Statement (on Open Scientific Research Data), <https://www.gov.uk/government/news/g8-science-ministers-statement>: s.n

G20 (2016). G20 Leaders' Communique Hangzhou Summit [Internet]. (6 de septiembre). Recuperado de http://www.consilium.europa.eu/en/meetings/international-summit/2016/09/Leaders-CommuniqueHangzhouSummit-final_pdf/

Naciones Unidas (2013). Report of the first session of the Plenary of the Intergovernmental Science-Policy Platform on Biodiversity, Bonn: PNUMA. Recuperado de http://www.ipbes.net/sites/default/files/downloads/IPBES_1_12_En.pdf.

Naciones Unidas, s.f. Decisión 14/86 del Consejo de Administración del PNUMA. s.l.:s.n.

Programa de las Naciones Unidas para el Medio Ambiente (2009). Gap analysis for the purpose of facilitating the discussions on how to improve and strengthen the Science-Policy Interface on Biodiversity and Ecosystem Services. s.l.: PNUMA.

Programa de las Naciones Unidas para el Medio Ambiente (2010). Busan Outcome. Busán: PNUMA Recuperado de <http://www.cid.org.nz/assets/Misc-Uploads/Misc-Reports/111122-Brief-on-Busan.pdf>).

Programa de las Naciones Unidas para el Medio Ambiente (2011). Report of the first session of the plenary meeting to determine modalities and institutional arrangements for an intergovernmental science-policy platform on biodiversity and ecosystem services., Nairobi: Recuperado de http://www.ipbes.net/sites/default/files/downloads/UNEP_IPBES_MI_1_8_EN.pdf.

Programa de las Naciones Unidas para el Medio Ambiente (2012). Report of the second session of the plenary meeting to determine modalities and institutional arrangements for an intergovernmental science-policy platform on biodiversity and ecosystem services, Panamá. Recuperado de http://www.ipbes.net/sites/default/files/downloads/UNEP_IPBES_MI_2_9_EN_0.pdf.

Programa de las Naciones Unidas para el Medio Ambiente (2015). Report of the Plenary of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on the work of its third session, Bonn. Recuperado de http://www.ipbes.net/sites/default/files/downloads/IPBES_3_18_EN.pdf.

Índice de figuras

Figura 1.	Tiempo geológico en millones de años	37
Figura 2.	Cuadro sintético de los órganos del CDB. Fuente: Alonso García, 2008.	40
Figura 3.	Infraestructuras de CREaTIVE-B. Fuente: Alonso García, 2014.	46
Figura 4.	Evolución de la red telefónica. Fuente: Ayala Mojica, s.f.	52
Figura 5.	Principio de computación de paquetes. Fuente: Cardozo, s. f.	53
Figura 6.	Mapa geográfico de ARPANET de septiembre de 1971. Fuente: Fernández Muñoz, s. f.	54
Figura 7.	Mapa geográfico de ARPANET de octubre 1980. Fuente: Fernández Muñoz, s. f.	54
Figura 8.	Esquema conceptual: revolución digital. Fuente: Martínez Navaro y Turégano García.	60
Figura 9.	Crecimiento de la información y tendencias del almacenamiento (proyecciones de crecimiento de la tasa de crecimiento de almacenamiento disponible de creación de información global). Fuente: IDC Digital Universe White Paper. Fuente: Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010.	64
Figura 10.	Data Warehouse. Fuente: Wikipedia.	65
Figura 11.	La ciberinfraestructura como una distribución entre los ejes técnico/social y global/local. Fuente: Knobel, 2007.	69
Figura 12.	A menudo se ha utilizado un modelo en capas para describir la «infraestructura de la información». En origen, la infraestructura de información (que recibe distintos nombres) incluye el hardware informático, las redes y las telecomunicaciones, tarjetas de memoria y el software operativo para hacer que estos trabajen conjuntamente. En esta concepción geológica de recursos interconectados de información global y telecomunicaciones, las capas inferiores de los sistemas informáticos y las redes permiten las capas superiores de middleware, colecciones de contenido digital y aplicaciones. Esta representación gráfica sugiere que el contenido digital en sí mismo puede y debe considerarse una infraestructura de estatus comparable o superior al otorgado al hardware y redes de computación. Fuente: Griffith, 2005.	70
Figura 13.	Paradigmas de la ciencia. Fuente: Hey <i>et al.</i> , 2009.	74
Figura 14.	Gestión de los ciclos de vida de los datos científicos en la e-Ciencia. Fuente: Demchenko <i>et al.</i> , 2013.	76
Figura 15.	Arquitectura de modelo propuesta. Fuente: Demchenko <i>et al.</i> , 2013.	77

Figura 16.	De los flujos de trabajo científicos a la infraestructura basada en la nube. Fuente: Demchenko <i>et al.</i> , 2013.	78
Figura 17.	Una e-infraestructura europea para y de repositorios digitales de e-ciencia. Fuente: Comunidades Europeas, 2008.	80
Figura 18.	ESFRI Roadmap 2010.	82
Figura 19.	Dibujo de Auke Herrema.	91
Figura 20.	Visión convencional de la jerarquía del conocimiento. Fuente: Tuomi, 1999.	92
Figura 21.	La jerarquía inversa. Fuente: Tuomi, 1999.	96
Figura 22.	Ciclo de datos tradicional. Fuente: Conner et al.	97
Figura 23.	Ciclo de vida de los datos extendido. Fuente: Conner et al.	98
Figura 24.	Dibujo de Auke Herrema.	99
Figura 25.	Posibles principios para una conceptualización de limitaciones y excepciones. Fuente: Gervais, 2008.	133
Figura 26.	El paraguas de la Open Science. Fuente Grigorov.	136
Figura 27.	Imagen original via doi: 10.1038/461145a. La investigación no puede florecer sin preservar los datos y hacerlos accesibles. La gestión de datos es necesaria en cada entramado de la ciencia.	144
Figura 28.	Crecimiento de los mandatos de acceso abierto (basado en el roadmap desde el 25 de mayo de 2012). Fuente: ROADMAP Schmidt y Kuchma, 2012.	149
Figura 29.	Los cinco elementos de INSPIRE. Fuente: Bartha y Kocsis, 2011.	156
Figura 30.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	158
Figura 31.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	159
Figura 32.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	160
Figura 33.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	161
Figura 34.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	161
Figura 35.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2017.	162
Figura 36.	Datos de investigación en abierto en Horizonte 2020. Fuente: Comisión Europea, 2016.	163
Figura 37.	Fuente: N. Kroes.	165
Figura 38.	Notes from a roundtable on open data at the White House. Fuente: sunlightfoundation.com/2017/07/26/white-house-open-data-roundtable .	175
Figura 39.	Basado en el Plan de Implementación de GEOSS, esquema de la transformación de las observaciones en información predictiva y de situación actual. En la parte derecha: productos de un sistema de pronóstico de sistemas terrestres y modelos especializados asociados organizados en las categorías de beneficios socioeconómicos de GEOSS, estratificados por el tiempo de entrega en el que requieren los	

	entregables (evaluaciones de las situaciones actuales, estimaciones de intervalos temporales, estudios a largo plazo de reanálisis). En la parte izquierda: requisitos de observación para un modelo de sistema terrestre integral, que además de los datos in situ incluye más datos satelitales actuales y estimados (satélites operacionales y de investigación). Las misiones operacionales y de investigación actuales cumplen muchos de los requerimientos espaciales, algunos de los cuales todavía no se cumplen. Las misiones satelitales se clasifican en función del ámbito geofísico observado (atmósfera, océano, superficie terrestre, tierra sólida) y por la frecuencia electromagnética utilizada por las misiones. En el centro: las herramientas que requieren un centro de pronóstico de sistemas terrestres (por ejemplo, ECMWF) y de sitios especializados de clientes/sitios asociados, con la finalidad de transformar las observaciones en información. Fuente: Hollingsworth <i>et al.</i> , 2005.	189
Figura 40.	Representación gráfica. Fuente: Nativi, Mazzetti, Santoro, Papeschi y Craglia, 2015.	190
Figura 41.	Imagen extraída de la página principal de GEOSS. Fuente: http://www.earthobservations.org/geoss.php .	193
Figura 42.	Arquitectura conceitual de alto nivel para la plataforma SW4SW. Fuente: Moodley <i>et al.</i> , 2008.	195
Figura 43.	Estrategia de Data Management Principles. Fuente: Alonso García, 2016a.	202
Figura 44.	Representación gráfica. Fuente: Hodson, 2015.	206
Figura 45.	Distintos niveles de interoperabilidad. Fuente: Van der Veer, 2008.	206
Figura 46.	Representación gráfica. Fuente: Landsbergen y Wolken, 1998.	209
Figura 47.	Políticas de datos para la participación pública en la investigación científica elaborado por Anne Bowser. Fuente: Alonso García, 2016b.	215
Figura 48.	Resumen de los mecanismos legales que promueven el acceso abierto y la interoperabilidad legal. Fuente: Doldirina, Eisentadt, Onsrud y Uhlir (pendiente de publicación).	223
Figura 49.	Cómo aplicar los principios FAIR. Fuente: Wilkinson <i>et al.</i> , 2016.	233
Figura 50.	Distintos niveles de FAIRness. Cuantos más elementos estén coloreados más FAIR son. Por ejemplo, si se añade un identificador persistente mejora el nivel de FAIRness de ese elemento. Los elementos en verde representan elementos FAIR y open, los elementos en rojo FAIR y closed	235
Figura 51.	Conjuntos estructurales de conjuntos de datos científicos. Fuente: Contreras y Reichman, 2015.	239
Figura 52.	Estructura de los órganos de gobierno del CDB. Fuente: http://www.cbd.int/cop .	251
Figura 53.	Procesos interseccionales del SBSSTA. Fuente: Koezt <i>et al.</i> , 2008.	256
Figura 54.	La complejidad de la biodiversidad. Fuente: Morin, Louafi, Orsini y Oubenal, 2018.	260
Figura 55.	Interacciones entre los Aichi Targets. A la izquierda, pueden verse las interacciones netas de los distintos objetivos (en barras) medidos respecto a la diferencia entre la suma de las interacciones descendentes que es igual a la suma de los impactos que afectan a los demás objetivos,	

y la suma de las interacciones ascendentes, que es igual a la suma de los impactos recibidos de los demás objetivos. Las acciones dirigidas a los impactos sobre los objetivos.

261

Figura 56. Políticas de indicadores de ciclos. Marco para usar indicadores para informar la política: la evaluación del problema implica definir objetivos generales de política, desarrollar modelos del sistema, seleccionar indicadores que reflejen los cambios de interés, definir objetivos específicos que puedan medirse con los indicadores y definir un conjunto de acciones o políticas para lograr los objetivos, evaluados con los modelos. Tras la evaluación, se eligen e implementan acciones, con el consiguiente cambio en la biodiversidad y en los factores de pérdida, que también pueden verse afectados por otros factores externos no relacionados con las acciones; el cambio se supervisa con los indicadores y se evalúa en relación con los objetivos, y ello lleva a una reevaluación. Las principales fuentes de incertidumbre y posibles fallas a lo largo del ciclo se numeran y discuten en detalle en la tabla S1: 1) supuestos en el proceso de evaluación; 2) vínculo entre evaluación y selección de acciones; 3) vínculo entre la selección y la implementación de acciones; 4) impacto de la acción que difiere del impacto anticipado; 5) vínculo entre cambio de biodiversidad y cambio de indicador; 6) vínculo entre el cambio de indicador y la evaluación del objetivo y 7) desajustes en escalas temporales y espaciales a lo largo del ciclo. Fuente: Nicholson *et al.* 2012.

263

Figura 57. Opinión actual en lo que respecta a la sostenibilidad medioambiental. Fuente: Díaz *et al.*, 2015.

270

Figura 58. Desajuste y coincidencia de instituciones de interfaz ciencia-política colaborativas frente a las lineales. Fuente: Koetz y Bridgewater, 2012.

271

Figura 59. Ilustración de las mayores tendencias en biodiversidad informática: hasta la fecha la disponibilidad de los datos y de la tecnología han sido los motores de muchas de las ideas y de los conceptos en este ámbito. Lo deseable sería que tanto las ideas como los conceptos sean los motores que desarrollen la nueva tecnología, los nuevos datos y los nuevos recursos de datos. Fuente: Peterson *et al.*, 2010.

280

Figura 60. Resumen en los términos más amplios de la biodiversidad informática a nivel mundial. Las cajas azules son los elementos básicos que subyacen en los procesos biológicos, desde el genotipo, fenotipo hasta la ecología y la pérdida de biodiversidad. Las cajas amarillas representan los productos de la información de la biodiversidad que con frecuencia se exploran. Las etiquetas que se encuentran fuera de las cajas son ejemplos de recursos informativos o iniciativas para la mayor parte de elementos del diagrama. Fuente: Peterson *et al.*, 2010.

281

Figura 61. Representación de los países. Fuente: GBIF.

286

Figura 62. Una posible aproximación a la representación de las funciones del IPBES. Fuente: GBIF, 2012.

289

Figura 63.	Distribución de los datos a los que se puede acceder a través del portal de datos de GBIF. Fuente: GBIF, 2012.	290
Figura 64.	Reflejo de las limitaciones que existen. Fuente: Chavan e Ingwersen, 2009.	292
Figura 65.	Catálogo electrónico de los nombres de organismos conocidos que servirán de vínculo entre muchos de los tipos de bases de datos biológicos y no biológicos posibles. Esos vínculos permitirán una minería de datos inimaginable hasta ahora porque, en la actualidad, es difícil, si no imposible, descubrir correlaciones entre los distintos conjuntos de datos. Fuente: Edwards et al., 2000.	294
Figura 66.	GBIF, en la actualidad, se está centrando en los datos primarios de biodiversidad, en los que se encuentran en la base de la pirámide de información. Fuente: Saarenmaa, 2005.	294
Figura 67.	Componentes de la arquitectura del sistema de información de GBIF. Fuente: Saarenmaa, 2005.	295
Figura 68.	Resumen del conocimiento digitalmente accesible por países mundialmente, dibujado por la Global Biodiversity Information Facility en enero de 2014, que muestra los 10 registros que presentan los países versus lo que proporcionan las instituciones de otros países. Los países (muchos de distintos continentes), que no dan datos, no aparecen en el gráfico. La línea de puntos indica equilibrio entre los datos entregados y los recibidos por el país. Fuente: Peterson y Soberón, 2015.	298
Figura 69.	Cinco puntos de cultura. Fuente: GBIF, 2015.	300
Figura 70.	Cinco puntos relacionados de datos. Fuente: GBIF, 2015.	302
Figura 71.	Cinco puntos relacionados con las muestras. Fuente: GBIF, 2015.	303
Figura 72.	Cinco puntos relacionados con la capacidad de comprender. Fuente: GBIF, 2015.	304
Figura 73.	Partiendo tanto de las observaciones de la tierra y en remoto como de los indicadores medioambientales, es necesario recoger los datos, dejarlos listos para su uso conjunto e integrarlos para preparar los indicadores. Fuente: Dubois <i>et al.</i> , 2010.	309
Figura 74.	Muestra de la complejidad existente en Europa. Fuente: Wetzel <i>et al.</i> , 2015.	316
Figura 75.	El EU BON y el desarrollo de la visión del GEO BON. Fuente: Hoffmann <i>et al.</i> , 2014.	318
Figura 76.	Actividades principales realizadas en las distintas capas de la e-infraestructura de LifeWatch. Fuente: LifeWatch, 2015.	322
Figura 77.	Impacto antropológico. Fuente: LifeWatch, 2015.	324
Figura 78.	Ejemplos de los principales servicios que proporcionará la plataforma LifeWatch. Fuente: LifeWatch, 2015.	326
Figura 79.	Relación de LifeWatch con otras infraestructuras e iniciativas pan-europeas. Fuente: LifeWatch, 2015.	327
Figura 80.	Resumen de los elementos de la ciberinfraestructura de DataONE que ilustran el valor que aporta a los usuarios asociados a un nodo miembro de biodiversidad. Colectores de datos de biodiversidad (a) depositar los	

	datos en un repositorio de biodiversidad miembro de DataONE. Usuarios de datos de biodiversidad (b) pueden acceder a dichos datos directamente a través del repositorio de biodiversidad. La media luna alrededor del repositorio representa la pila de software que permite a los nodos miembros que el repositorio funcione. Esta pila de software la instalan los responsables de la ciberinfraestructura de DataONE. Los usuarios de biodiversidad pueden seguir accediendo a los datos tal como lo hacían antes (b); sin embargo, a través de la vinculación a la ciberinfraestructura de DataONE (c), dichos datos están accesibles a una comunidad más amplia de usuarios de DataONE (d) a través de la búsqueda y recuperación habilitados. Fuente: Michener et al., 2012.	329
Figura 81.	Relaciones entre las comunidades de stakeholders: los científicos son los stakeholders primarios y los círculos representan los entornos de investigación científica, los stakeholders secundarios están asociados con cada entorno de investigación científica e incluye organizaciones (cajas) e individuos(figuras ovaladas), las cajas con líneas discontinuas se refieren a los stakeholders asociados con cada nivel de gobierno. Fuente: Michener et al., 2012.	330
Figura 82.	Organización de DataONE. Fuente: DataONE, s. f.	331
Figura 83.	Estructura de DataONE. Fuente: DataONE, s. f..	332
Figura 84.	DataONE Membership Model. Fuente: Adaptada de DataONE, s.f.	333
Figura 85.	Neon Climate Domains. Fuente: NEON, 2006.	338
Figura 86.	NEON difiere de otras redes de monitorización medioambiental porque, en su diseño, integra procesos, interacciones y respuestas. Fuente: Keller et al., 2008.	339
Figura 87.	Estructura SANBI. Fuente: Alonso García, 2014.	342
Figura 88.	Trabajo de Sanbi. Fuente: Sanbi, 2016.	343
Figura 89.	Implementación de la infraestructura de ALA. Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.	347
Figura 90.	Infraestructura de ALA. Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.	348
Figura 91.	Modelo de arquitectura de ALA. Fuente: Fuente: Atlas of Living Australia sharing biodiversity knowledge, 2016.	349
Figura 92.	Como interactúan los distintos instrumentos. Fuente: Anne Larigauderie.	352
Figura 93.	Ejemplos de movilización de datos, en este caso del EU BON. Refleja muy bien la interrelación entre las diferentes herramientas existentes. Fuente: Wetzel et al., 2015.	353
Figura 94.	How we fit together: global view. Fuente: CREaTIVE-B.	358
Figura 95.	Rol de las e-infraestructuras en la facilitación de la colaboración. Fuente: Alonso García et al., 2014.	359
Figura 96.	Evaluación del potencial de lograr la interoperabilidad. Visión general. Fuente: Alonso García et al., 2014.	360
Figura 97.	Evaluación del potencial de lograr la interoperabilidad. Lógica de los servicios. Fuente: Alonso García et al., 2014.	361

Figura 98.	Evaluación del potencial de lograr la interoperabilidad. Datos. Fuente: Alonso García et al., 2014.	361
Figura 99.	Posibles pasos para el cálculo de las Variables Esenciales de Biodiversidad (EBV, verde) y relacionadas preguntas y retos científicos (azul) y técnicos (rojo). Fuente: Konijn, 2016.	363
Figura 100.	Perspectiva de los 4 talleres de GLOBIS-B relacionados con distribución de especies y abundancia (1 y 2), características de especies (3) e interacciones de especies (4). Fuente: Konijn, 2016.	364
Figura 101.	Presentación gráfica de los elementos del plan de trabajo de GLOBIS-B integrado por 5 paquetes de trabajo. Los paquetes de trabajo WP2-4 se encargan del contenido de los workshops mientras que los paquetes WP1 y WP5 son, fundamentalmente de carácter administrativo	365
Figura 102.	Caracterización de las infraestructuras de investigación que cooperan en el proyecto GLOBIS-B. Las infraestructuras de investigación son distintas en su ámbito geográfico (eje-y) y su foco en movilización de datos vs. producción de conocimiento eje-x. Esta imagen se ha hecho tomando como base otra realizada para el proyecto CREaTIVE-B. Fuente: Kissling et al. 2015.	366
Figura 103.	Camino propuesto para pasar de la observación de la tierra a las EBVs, a los indicadores de biodiversidad y a los Aichi Targets con ejemplos de observación de la tierra para hacer mediciones parciales o completas de las EBVS. Fuente: O'Connor et al. 2015.	367
Figura 104.	GEO BON. Fuente GEO BON. Fuente: GEO BON.	368
Figura 105.	Variables Esenciales de Biodiversidad. Fuente: Pereira, 2013.	369
Figura 106.	Diagrama conceptual que ilustra la intensidad de la pérdida y alteración asociadas con las distintas dimensiones del cambio de la biodiversidad: extinciones, pérdida de diversidad genética, cambios en la abundancia y estructura de la comunidad y zonas de ajuste. Fuente: Pereira et al., 2012.	370
Figura 107.	Índice de poblaciones de vida en el planeta. Fuente: Pereira, 2013.	371
Figura 108.	E-infraestructuras de investigación especializadas en datos de biodiversidad. Fuente: Alonso, 2016a.	373
Figura 109.	Proceso para que las variables esenciales de biodiversidad sean operativas. Fuente: Schmeller et al., 2017.	374
Figura 110.	Ejemplos de EBV potenciales, representación en el espacio durante un tiempo determinado y forma en la que pueden usarse para documentar cambios temporales de biodiversidad a la hora de comparar distintas representaciones espaciales en representaciones temporales. Fuente: Schmeller et al., 2017.	375
Figura 111.	Las EBV forman parte de la cadena de información. Están conceptualmente colocadas entre los datos en bruto y los indicadores. Fuente: Kissling et al. 2017.	376
Figura 112.	Estrategia de Data Management Principles. Fuente: Alonso García, 2016a.	382
Figura 113.	Políticas de datos para la participación pública en la investigación científica elaborado por Anne Bowser. Fuente: Alonso García, 2016a.	383

Figura 114.	Distintos niveles de interoperabilidad. Fuente: Van der Veer, 2008.	394
Figura 115.	Cómo aplicar los principios FAIR. Fuente: Wilkinson et al., 2016.	403
Figura 116.	Data management principles branding. Fuente: Halpern, 2015.	404